# Secure Collaborations for the Industrial Internet of Things

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Master of Science

**Jan Pennekamp**

aus Hilden

Berichter:

Prof. Dr.-Ing. Klaus Wehrle
Prof. Dr.-Ing. Florian Kerschbaum

Tag der mündlichen Prüfung: 22.09.2023

**Reports on Communications and Distributed Systems**

edited by
Prof. Dr.-Ing. Klaus Wehrle
Communication and Distributed Systems,
RWTH Aachen University

Volume 23

**Jan Pennekamp**

# Secure Collaborations for the Industrial Internet of Things

Shaker Verlag
Düren 2024

**Bibliographic information published by the Deutsche Nationalbibliothek**
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: D 82 (Diss. RWTH Aachen University, 2023)



Printed in Germany.

ISBN 978-3-8440-9467-1
ISSN 2191-0863

Shaker Verlag GmbH • Am Langen Graben 15a • 52353 Düren
Phone: 0049/2421/99011-0 • Telefax: 0049/2421/99011-9
Internet: www.shaker.de • e-mail: info@shaker.de

## Abstract

The Industrial Internet of Things (IIoT) is leading to increasingly-interconnected and networked industrial processes and environments, which, in turn, results in stakeholders gathering vast amounts of information. Although the global sharing of information and *industrial collaborations* in the IIoT promise to enhance productivity, sustainability, and product quality, among other benefits, most information is still commonly encapsulated in local information silos. In addition to interoperability issues, *confidentiality concerns* of involved stakeholders remain the main obstacle to fully realizing these improvements in practice as they largely hinder real-world industrial collaborations today. Therefore, this dissertation addresses this mission-critical research gap. Since existing approaches to privacy-preserving information sharing are not scalable to industry-sized applications in the IIoT, we present solutions that enable *secure* collaborations in the IIoT while providing technical (confidentiality) guarantees to the involved stakeholders. Our research is crucial (i) for demonstrating the potential and added value of (secure) collaborations and (ii) for convincing cautious stakeholders of the usefulness and benefits of technical building blocks, enabling reliable sharing of confidential information, even among direct competitors.

Our *interdisciplinary* research thus focuses on establishing and realizing secure industrial collaborations in the IIoT. In this regard, we study two overarching angles of collaborations in detail. First, we distinguish between collaborations along and across supply chains, with the former type entailing more relaxed confidentiality requirements. Second, whether or not collaborators know each other in advance implies different levels of trust and requires different technical guarantees. We rely on well-established building blocks from *private computing* (i.e., privacy-preserving computation and confidential computing) to reliably realize secure collaborations. We thoroughly evaluate each of our designs, using multiple real-world use cases from production technology, to prove their practical feasibility for the IIoT.

By applying private computing, we are indeed able to secure collaborations that not only scale to industry-sized applications but also allow for use case-specific configurations of confidentiality guarantees. In this dissertation, we use well-established building blocks to assemble novel solutions with technical guarantees for all types of collaborations (along and across supply chains as well as with known or unknown collaborators). Finally, on the basis of our experience with engineers, we have derived a *research methodology* for future use that structures the process of interdisciplinary development and evaluation of secure collaborations in the evolving IIoT.

Overall, given the aforementioned improvements, our research should greatly contribute to convincing even cautious stakeholders to participate in (reliably-secured) industrial collaborations. Our work is an essential first step toward establishing widespread information sharing among stakeholders in the IIoT. We further conclude: (i) collaborations can be reliably secured, and we can even provide technical guarantees while doing so; (ii) building blocks from private computing scale to industrial applications and satisfy the outlined confidentiality needs; (iii) improvements resulting from industrial collaborations are within reach, even when dealing with cautious stakeholders; and (iv) the interdisciplinary development of sophisticated yet appropriate designs for use case-driven secure collaborations can succeed in practice.

## Kurzfassung

Das industrielle Internet der Dinge (IIoT) führt zu vernetzten industriellen Prozessen, wodurch viele Informationen gesammelt werden. Obwohl der globale Austausch von Informationen und *industrielle Zusammenarbeit* erhebliche Verbesserungen (wie z.B. Produktivität, Nachhaltigkeit, Produktqualität und weiteres) versprechen, sind die Daten häufig nur lokal zugängig. Neben Interoperabilitätsproblemen behindern heutzutage vor allem *Vertraulichkeitsbedenken* die Etablierung von industrieller Zusammenarbeit. Mit dieser Dissertation adressieren wir diese Bedenken. Da bestehende Konzepte zum sicheren Teilen von Informationen nicht für industrielle Zwecke geeignet sind, stellen wir Lösungen vor, die eine *sichere* Zusammenarbeit im IIoT ermöglichen und gleichzeitig technische Garantien bieten. Unsere Forschung ist von entscheidender Bedeutung, um (i) das Potenzial und den Mehrwert von (sicherer) Zusammenarbeit aufzuzeigen und (ii) reservierte Unternehmen vom Nutzen und den Vorteilen technischer Bausteine zu überzeugen, die einen zuverlässigen Austausch vertraulicher Informationen ermöglichen, selbst zwischen direkten Wettbewerbern.

Unsere *interdisziplinäre* Forschung konzentriert sich daher auf die Etablierung und Realisierung von sicherer industrieller Zusammenarbeit im IIoT. Wir unterscheiden dabei nicht nur zwischen Kooperationen entlang und über Lieferketten hinweg, sondern auch, ob sich die beteiligten Unternehmen im Voraus kennen oder nicht. Diese Dimensionen zeigen verschiedene Vertrauensverhältnisse auf und benötigen somit in der Umsetzung unterschiedlich starke technische Garantien. Wir verwenden dabei bewährte technische Bausteine um vertrauenswürdige industrielle Zusammenarbeiten zuverlässig zu realisieren. Wir evaluieren unsere vorgestellten Entwürfe umfangreich anhand von realen Anwendungsfällen aus dem Bereich der Produktionstechnik, auch um ihren praktischen Nutzen für Unternehmen im IIoT zu belegen.

Der Einsatz der bewährten Bausteine erlaubt uns in der Tat Lösungen zu erstellen, die nicht nur sicher sind, sondern auch für den Einsatz in verschiedenen industriellen Anwendungsszenarien geeignet sind. In dieser Dissertation haben wir etablierte Bausteine kombiniert um neuartige Lösungen mit technischen Garantien für alle Arten von industrieller Zusammenarbeit (entlang und über Lieferketten hinweg sowie mit bekannten oder unbekannten Unternehmen) zu realisieren. Basierend auf unseren Erfahrungen in der Zusammenarbeit mit Ingenieuren haben wir außerdem eine *Methodik für interdisziplinäre Forschung* hergeleitet, die diesen Prozess strukturiert.

Insgesamt sollten unsere Forschungsergebnisse angesichts der zu erwartenden Verbesserungen einen Beitrag leisten, auch zurückhaltende Unternehmen zu überzeugen, sich an (zuverlässig gesicherten) industriellen Kooperationen zu beteiligen. Unsere Arbeit ist somit ein wesentlicher Schritt zur Etablierung von industrieller Zusammenarbeit im IIoT. Außerdem folgern wir aus unseren Ergebnissen: (i) industrielle Kooperationen können zuverlässig abgesichert werden, und wir können dabei sogar technische Garantien bieten, (ii) bestehende Bausteine zum sicheren Informationsaustausch können auch im industriellen Kontext angewendet werden, (iii) Verbesserungen, die sich aus industrieller Zusammenarbeit ergeben, sind somit in Reichweite, selbst wenn man mit skeptischen Unternehmen zu tun hat, und (iv) die interdisziplinäre Entwicklung anspruchsvoller und dennoch geeigneter Designs für anwendungsbezogene sichere Zusammenarbeit kann auch für die praktische Nutzung gelingen.

# Acknowledgments

First and foremost, I want to express my gratitude to Andriy, Fabian, and Martin, who got me excited about research as part of their international collaboration during my Bachelor's degree. Without you, I would not be where I am today and I would not even have started this chapter. Moreover, concluding this chapter has only been possible since Prof. Klaus Wehrle and Prof. Florian Kerschbaum served as evaluators of this dissertation. I would like to particularly thank Klaus for allowing me to pursue my own ideas, doing my doctorate part-time, and giving me the freedom to develop myself in many ways and different research domains. Moreover, I am grateful to Florian for the pleasant exchange, despite initially approaching you as a greenhorn, as well as for going out of your way to take such a notable role in my doctoral committee. Lastly, as I would have expected following our previous encounters, Prof. Wil van der Aalst and Prof. Stefan Decker served with the best professional conduct on my committee, which is greatly appreciated. Thank you all!

Second, I am fortunate that throughout my doctoral studies, several gears engaged quite well: (i) Interdisciplinary research in the Cluster of Excellence "Internet of Production" exposed me to many new perspectives, colleagues, and ways of working. (ii) Joint research with (distinguished) students and graduate assistants frequently led to publication, also due to the support of various (interdisciplinary) co-advisors. (iii) Several colleagues (at COMSYS) greatly supported me with my paper writing, especially Roman, Markus, Lennart, Erik, and Martin. Likewise, Joscha provided me with on-point feedback while revising the numerous plots of this dissertation. Unfortunately, naming each and every one would go beyond the scope, but please be sure that I have neither forgotten your contributions nor what I have learned in every project, paper, and cooperation. In this context, I even had the pleasure of coordinating and agreeing on (larger) international research activities and projects. These elements essentially form the basis of this dissertation. (iv) Without the pandemic, my research would not have progressed as it did (a limited social life and the lack of business trips can do wonders…). Still, I had to take a sprint while writing up this dissertation (first as part of our Dagstuhl writing retreat and then in parallel to the daily madness at the office). (v) Last but not least, even though a significant part of this dissertation was created while working from home for over two years, knowing to have a good soul at the office has been a pleasure. Claudia, thank you for having an open ear and for taming the finances and all personnel matters. Combined, these gears and my CV of failures make up my academic journey so far.

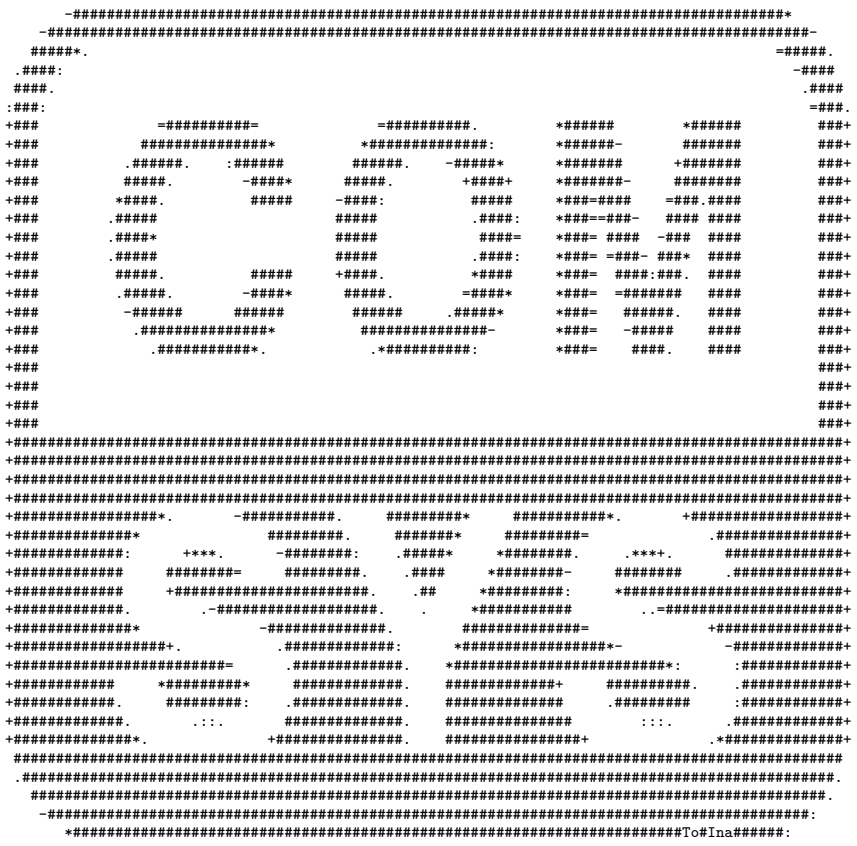
Third, my proofreaders (Eric, Erik, Ike, Ina, Jens, Johannes, Lennart, Markus, Martin, Robert, Roman, Rut, and the language center) might have overlooked some mistakes (just like me). Now, these issues will be retained for eternity; so, thanks!

Fourth, apart from these content-specific matters, I would also like to say thank you for the privilege of traveling together with various colleagues (Martin, Ike, Helge, Roman, Johannes, and Christian), as well as meeting friends and colleagues along these journeys (Philip, Andriy, Asya, Sebastian, Fritz, Florian—twice—, Dominik, Erik, Angelika, Tamer, David, and Lea). Additionally, I would not be the same without the role model function of Martin, Torsten, Jan, Jens, and Roman, as well

as the inventors of the conference call. Then again, I would also like to take this opportunity to thank Alex, Constantin, David and David, Dirk, Eric, Felix, Ike, Ina, Johannes, JT, Kai, Konrad, Lars, Liam, Niklas, Patrick, Philipp, Rainer, René, Salil, Sascha, Stefan, Timo, and Yannik. I believe you know your place in my story.

Fifth, primarily, Moritz and Malte made sure that I had distractions from work through our vacations; but this note also applies to so many others. Moreover, my friends helped me navigate a few challenging times in the past: Here, I would like to mention Marco first and foremost, but also Felix, Moritz, Robert, Roman, and Ina.

Finally, to my friends, brother, and colleagues (i.a., Erik, Erik, Felix, Julius, Malte ⓒ, Simon, and Tim) who are still pursuing their doctorates at this point, go on, and pull through. I looking forward to receiving invitations to attending your defenses down the road! You will be just fine and turn out to be worthy doctorate holders.




## Funding Acknowledgment

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

# Published Papers

Parts of this dissertation are based on the following peer-reviewed papers that have already been published. All my collaborators are among my co-authors. A detailed attribution of contributions can be found after the list of supervised theses.

## List of Publications


[BDJ+22] Philipp Brauner, Manuela Dalibor, Matthias Jarke, Ike Kunze, István Koren, Gerhard Lakemeyer, Martin Liebenberg, Judith Michael, **Jan Pennekamp**, Christoph Quix, Bernhard Rumpe, Wil van der Aalst, Klaus Wehrle, Andreas Wortmann, and Martina Ziefle. A Computer Science Perspective on Digital Transformation in Production. *ACM Transactions on Internet of Things*, 3(2), May 2022.

[BPM+21] Lennart Bader, **Jan Pennekamp**, Roman Matzutt, David Hedderich, Markus Kowalski, Volker Lücken, and Klaus Wehrle. Blockchain-Based Privacy Preservation for Supply Chains Supporting Lightweight Multi-Hop Information Accountability. *Information Processing & Management*, 58(3), May 2021.

[GPL+20] Lars Gleim, **Jan Pennekamp**, Martin Liebenberg, Melanie Buchsbaum, Philipp Niemietz, Simon Knape, Alexander Epple, Simon Storms, Daniel Trauth, Thomas Bergs, Christian Brecher, Stefan Decker, Gerhard Lakemeyer, and Klaus Wehrle. FactDAG: Formalizing Data Interoperability in an Internet of Production. *IEEE Internet of Things Journal*, 7(4):3243–3253, April 2020.

[PAB+24] **Jan Pennekamp**, Fritz Alder, Lennart Bader, Gianluca Scopelliti, Klaus Wehrle, and Jan Tobias Mühlberg. Securing Sensing in Supply Chains: Opportunities, Building Blocks, and Designs. *IEEE Access*, 12:9350–9368, January 2024.

[PAM+20] **Jan Pennekamp**, Fritz Alder, Roman Matzutt, Jan Tobias Mühlberg, Frank Piessens, and Klaus Wehrle. Secure End-to-End Sensing in Supply Chains. In *Proceedings of the 2020 IEEE Conference on Communications and Network Security (CNS '20)*. IEEE, July 2020. Proceedings of the 5th International Workshop on Cyber-Physical Systems Security (CPS-Sec '20).

[PBD+21] **Jan Pennekamp**, Erik Buchholz, Markus Dahlmanns, Ike Kunze, Stefan Braun, Eric Wagner, Matthias Brockmann, Klaus Wehrle, and Martin Henze. Collaboration is not Evil: A Systematic Look at Security Research for Industrial Use. In *Proceedings of the Workshop on Learning from Authoritative Security Experiment Results (LASER '20)*. ACSAC, December 2021.

[PBL+20] **Jan Pennekamp**, Erik Buchholz, Yannik Lockner, Markus Dahlmanns, Tiandong Xi, Marcel Fey, Christian Brecher, Christian Hopmann, and Klaus Wehrle. Privacy-Preserving Production Process Parameter Exchange. In *Proceedings of the 36th Annual Computer Security Applications Conference (ACSAC '20)*, pages 510–525. ACM, December 2020.

[PBM+20] **Jan Pennekamp**, Lennart Bader, Roman Matzutt, Philipp Niemietz, Daniel Trauth, Martin Henze, Thomas Bergs, and Klaus Wehrle. Private Multi-Hop Accountability for Supply Chains. In *Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops '20)*. IEEE, June 2020. Proceedings of the 1st Workshop on Blockchain for IoT and Cyber-Physical Systems (BIoTCPS '20).

[PDF+23] **Jan Pennekamp**, Markus Dahlmanns, Frederik Fuhrmann, Timo Heutmann, Alexander Kreppein, Dennis Grunert, Christoph Lange, Robert H. Schmitt, and Klaus Wehrle. Offering Two-Way Privacy for Evolved Purchase Inquiries. *ACM Transactions on Internet Technology*, 23(4), November 2023.

[PDG+19] **Jan Pennekamp**, Markus Dahlmanns, Lars Gleim, Stefan Decker, and Klaus Wehrle. Security Considerations for Collaborations in an Industrial IoT-based Lab of Labs. In *Proceedings of the 3rd IEEE Global Conference on Internet of Things (GCIoT '19)*. IEEE, December 2019.

[PFD+21] **Jan Pennekamp**, Frederik Fuhrmann, Markus Dahlmanns, Timo Heutmann, Alexander Kreppein, Dennis Grunert, Christoph Lange, Robert H. Schmitt, and Klaus Wehrle. Confidential Computing-Induced Privacy Benefits for the Bootstrapping of New Business Relationships. Technical Report RWTH-2021-09499, RWTH Aachen University, November 2021. Blitz Talk at the 2021 Cloud Computing Security Workshop (CCSW '21).

[PGH+19] **Jan Pennekamp**, René Glebke, Martin Henze, Tobias Meisen, Christoph Quix, Rihan Hai, Lars Gleim, Philipp Niemietz, Maximilian Rudack, Simon Knape, Alexander Epple, Daniel Trauth, Uwe Vroomen, Thomas Bergs, Christian Brecher, Andreas Bührig-Polaczek, Matthias Jarke, and Klaus Wehrle. Towards an Infrastructure Enabling the Internet of Production. In *Proceedings of the 2nd IEEE International Conference on Industrial Cyber Physical Systems (ICPS '19)*, pages 31–37. IEEE, May 2019.

[PHS+19] **Jan Pennekamp**, Martin Henze, Simo Schmidt, Philipp Niemietz, Marcel Fey, Daniel Trauth, Thomas Bergs, Christian Brecher, and Klaus Wehrle. Dataflow Challenges in an *Internet* of Production: A Security & Privacy Perspective. In *Proceedings of the ACM Workshop on Cyber-Physical Systems Security & Privacy (CPS-SPC '19)*, pages 27–38. ACM, November 2019.

[PHW21] **Jan Pennekamp**, Martin Henze, and Klaus Wehrle. Unlocking Secure Industrial Collaborations through Privacy-Preserving Computation. *ERCIM News*, 126:24–25, July 2021. *Non-peer-reviewed paper.*

[PLV+23] **Jan Pennekamp**, Johannes Lohmöller, Eduard Vlad, Joscha Loos, Niklas Rodemann, Patrick Sapel, Ina Berenice Fink, Setz Schmitz, Christian Hopmann, Matthias Jarke, Günther Schuh, Klaus Wehrle, and Martin Henze. Designing Secure and Privacy-Preserving Information Systems for Industry Benchmarking. In *Proceedings of the 35th International Conference on Advanced Information Systems Engineering (CAiSE '23)*, pages 489–505. Springer, June 2023.

[PMK+21] **Jan Pennekamp**, Roman Matzutt, Salil S. Kanhere, Jens Hiller, and Klaus Wehrle. The Road to Accountable and Dependable Manufacturing. *Automation*, 2(3):202–219, September 2021.

[PMK+24] **Jan Pennekamp**, Roman Matzutt, Christopher Klinkmüller, Lennart Bader, Martin Serror, Eric Wagner, Sidra Malik, Maria Spiß, Jessica Rahn, Tan Gürpinar, Eduard Vlad, Sander J. J. Leemans, Salil S. Kanhere, Volker Stich, and Klaus Wehrle. An Interdisciplinary Survey on Information Flows in Supply Chains. *ACM Computing Surveys*, 56(2), February 2024.

[PSF+20] **Jan Pennekamp**, Patrick Sapel, Ina Berenice Fink, Simon Wagner, Sebastian Reuter, Christian Hopmann, Klaus Wehrle, and Martin Henze. Revisiting the Privacy Needs of Real-World Applicable Company Benchmarking. In *Proceedings of the 8th Workshop on Encrypted Computing & Applied Homomorphic Cryptography (WAHC '20)*, pages 31–44. HomomorphicEncryption.org, December 2020.

### Dissertation Digest

This dissertation has further been summarized as part of a peer-reviewed digest.

[Pen24] **Jan Pennekamp**. Evolving the Industrial Internet of Things: The Advent of Secure Collaborations. In *Proceedings of the 2024 IEEE/IFIP Network Operations and Management Symposium (NOMS '24)*. IEEE, May 2024.

## Supervised Theses

Parts of this dissertation are based on the following Bachelor's and Master's theses that were written by students under my supervision. For all theses, I developed the project topic, goals, and fundamental architectural and evaluation design, while my co-advisors provided feedback and domain expertise. The student respectively developed the detailed thesis concept, implementation, and evaluation, as well as the written thesis. All (co-)advisors contributed feedback and reviewed the theses.

## List of Supervised Theses

[Bad20] Lennart Bader. Privacy and Transparency in Digital Supply Chains. Advisors: **Jan Pennekamp**, Roman Matzutt, and Philipp Niemietz. Examiners: Klaus Wehrle and Thomas Bergs. Master's Thesis. RWTH Aachen University, February 2020.

[Buc20] Erik Buchholz. Privacy-Preserving Exchange of Process Parameters. Advisors: **Jan Pennekamp** and Yannik Lockner. Examiners: Klaus Wehrle and Christian Hopmann. Master's Thesis. RWTH Aachen University, June 2020.

[Fuh21] Frederik Fuhrmann. Two-way Privacy For Purchase Inquiries in Industry. Advisors: **Jan Pennekamp** and Timo Heutmann. Examiners: Klaus Wehrle and Robert H. Schmitt. Master's Thesis. RWTH Aachen University, January 2021.

[Jes21] Fabian Thorsten Jess. Enhancing Supply Chain Management with Trustworthy and Reliable Sensor Data. Advisors: **Jan Pennekamp**, Lennart Bader, and Fritz Alder. Examiners: Klaus Wehrle and Frank Piessens. Bachelor's Thesis. RWTH Aachen University, August 2021.

[Mic21] Jan-Gustav Michnia. Improving Privacy-Preserving Company Benchmarking with Modern FHE Schemes. Advisors: **Jan Pennekamp**, Niklas Rodemann, and Martin Henze. Examiners: Klaus Wehrle and Günther Schuh. Bachelor's Thesis. RWTH Aachen University, November 2021.

[Siu20] Alexander Stanislaw David Siuda. Web-Based Privacy-Preserving Comparison of KPIs. Advisors: **Jan Pennekamp** and Martin Henze. Examiners: Klaus Wehrle and Elmar Padilla. Bachelor's Thesis. RWTH Aachen University, October 2020.

[Vla22] Eduard Vlad. Applying Trusted Execution for Privacy-Preserving Company Benchmarking. Advisors: Johannes Lohmöller, **Jan Pennekamp**, and Patrick Sapel. Examiners: Klaus Wehrle and Christian Hopmann. Bachelor's Thesis. RWTH Aachen University, September 2022.

[Wag20] Simon Wagner. Privacy-Preserving Company Benchmarking. Advisors: **Jan Pennekamp**, Martin Henze, and Patrick Sapel. Examiners: Klaus Wehrle and Christian Hopmann. Bachelor's Thesis. RWTH Aachen University, September 2020.

# Attribution of Contributions

Parts of this dissertation are based on collaborations with students, researchers, and practitioners from industry. The resulting publications form the scientific foundation of this dissertation and were created with the support of the respective co-authors. We now attribute the parts of this dissertation to the respective publications, theses, and authors. We disseminate only parts of the attributed publications. Accordingly, these referenced publications contain additional information, evaluations, and discussions. If not explicitly stated otherwise, in addition to bootstrapping the collaborations with domain experts, the author of this dissertation was responsible for their initial concepts, methodologies, and designs, as well as the final publication.

**Background** M.H. initially suggested to pursue the research questions of two papers [PGH+19, PHS+19]. J.P. organized and managed the collaborations. J.P. and M.H. jointly developed the outline [PGH+19] and jointly worked on the content and presentation [PHS+19], which is based on discussions with P.N. and S.S. With input from P.N., R.G. contributed the fine blanking use case. M.R. and S.K. prepared the high-pressure die casting and connected job shop use cases, respectively. T.M., R.H., L.G., and C.Q. provided the expertise and presentation on data management. J.P., M.D., and L.G. jointly conducted the work [PDG+19] based on the concept of J.P. Furthermore, J.P. supported L.G. with his paper [GPL+20] by contributing his view on data security. Overall, L.G., J.P., M.L., M.B., P.N., and S.K. jointly worked on the paper. J.P., R.M., and J.H. jointly wrote the paper [PMK+21] following the initiative of J.P. The discussions with S.S.K. greatly supported our work.

**Use Cases** The use case descriptions are based on information (and visualizations) provided by domain experts. Their input is distributed to the selected applications as follows. P.N. is the fine-blanking expert [PGH+19, PHS+19, PBM+20]. D.H. provided the details of the urban electric vehicle [BPM+21]. While S.K. and T.X. assisted with the connected job shop [GPL+20, PBL+20], T.H., A.K., and D.G. conveyed their knowledge of procurement processes [PDF+23]. P.S. (injection molding) and N.R. (global production networks) supported the company benchmarking descriptions [PLV+23]. Y.L. contributed his injection molding expertise [PBL+20].

**Information Processing in Supply Chains** Our interdisciplinary survey on information flows in supply chains [PMK+24] is a collaborative effort that greatly exceeds the scope of this thesis. The relevant attribution of contributions for this dissertation is as follows. M.S., J.R., and T.G. drafted the general use cases in supply chains. Together with J.P. and C.K., these use cases were structured and repeatedly revised. In addition, in the context of our sensing paper [PAB+24], J.P. discussed the common sensing applications in supply chains (Section 4.1.2.2) after drafting them with M.S.

**Secure and Reliable (End-to-End) Sensing** J.P., R.M., and F.A. jointly developed the initial concept [PAM+20]. F.A. and J.T.M. contributed their expertise in trusted computing. J.P., F.A., and J.T.M. jointly worked on the cost evaluation and security discussion. Subsequently, based on a design by L.B. and J.P., F.T.J. implemented a proof of concept for his thesis [Jes21] to link the sensing part of the data processing pipeline with the information-sharing part. J.P., F.A., L.B., and J.T.M. further jointly evolved the initial conceptual design [PAB+24]. While G.S. conducted the evaluation of the sensing part of the data processing pipeline, L.B. provided the results of the blockchain evaluation.

**Long-Term Private (Multi-Hop) Information Sharing** Jointly, J.P. and R.M. proposed the initial concept [PBM+20]. L.B. significantly evolved this concept during his thesis [Bad20] and evaluated his implementation based on use case data provided by P.N. Aside from that, L.B. further contributed the concept's security discussion as part of his thesis, which was subsequently discussed in detail with J.P. and R.M. before inclusion in the publications [PBM+20, BPM+21]. J.P. discovered and secured (jointly with D.H.) the electric vehicle use case data, which was evaluated and described by L.B. [BPM+21].

**Finding New Suppliers with Privacy-Preserving Purchase Inquiries** For the thesis of F.F. [Fuh21], T.H. commented on the initial ideas of J.P. and provided the evaluation data. F.F. proposed to also pursue an approach in the direction of cHPI. J.P. and F.F. jointly evolved this initial idea and also came up with HPI. F.F. further implemented and evaluated the designs. For our joint paper [PDF+23], C.L. supported the author of this thesis with his knowledge of data modeling. Additionally, A.K. and D.G. contributed the use case description and helped clarify the steps in today's procurement processes.

**Privacy-Preserving Company Benchmarking** J.P. and M.H. discussed the initial concept [PSF+20], which S.W. implemented in his thesis [Wag20]. Based on use case data provided by P.S., S.R. conducted the first evaluation of PCB (the initial predecessor of SW-PCB). I.B.F. assisted with the presentation of the design and evaluation. Furthermore, A.S.D.S. implemented the WebAssembly-based client for his thesis [Siu20]. Subsequently, J.G.M. implemented a CONCRETE-based prototype of SW-PCB during his thesis [Mic21], which J.Loo. re-implemented with an updated library version. N.R. contributed another set of use case data. Based on the idea by J.P., J.Loh. and J.P. jointly advised E.V., who implemented HW-PCB, re-implemented SW-PCB in his thesis [Vla22], and evaluated them using both use cases. I.B.F., J.P., J.Loh., M.H., and M.J. discussed the structure of the follow-up paper [PLV+23].

**Privacy-Preserving Parameter Exchange** J.P. and Y.L. jointly identified the need for research [PBL+20]. J.P. proposed an initial design that was significantly evolved and improved in collaboration with E.B. during his thesis [Buc20]. Based on use case data provided by Y.L., E.B. extensively evaluated his implementation. J.P. and E.B. jointly identified another use case, which E.B. evaluated based on the data provided by T.X. While J.P. and M.D. worked on the line of presentation, E.B. greatly assisted during the writing process.

**Appraisal on Secure Industrial Collaborations** The outlook of this chapter is influenced by three publications [PHW21, PMK+21, BDJ+22], which J.P. proposed and initiated. Specifically, the strategic research directions are of interest [BDJ+22], which J.P. outlined, drafted in full, and edited afterward based on comments and suggestions by all co-authors. Moreover, the excursus is mainly based on our post-workshop paper [PBD+21]. J.P. initially suggested the presented process cycle, which M.H. later commented on. M.H., M.D., I.K., and E.W. further shared their experience. E.B. reported in detail on experience made during his thesis [Buc20] and our paper submission [PSF+20]. I.K. prepared the discussion on related work after discussions with J.P. Furthermore, S.B. contributed insights on RDM. M.B. sketched a real-world example and commented on our work in light of his domain expertise.

# Contents

# 1

# Introduction

In this dissertation, we focus on different types of collaborations in the Industrial Internet of Things (IIoT). Primarily, we consider the information security dimension as it is crucial to address the confidentiality requirements of involved stakeholders, which currently hinders collaborations from being applied at broad scale. Thus, we detail how to implement collaborations securely using (well-established) technical building blocks that effectively entail reliability and confidentiality guarantees.

We start with the motivation for and a broad introduction to collaborations in the first chapter. In particular, in Section 1.1, we motivate our research direction before specifically introducing collaborations in the IIoT and the associated security and privacy challenges in Section 1.2. Finally, in Section 1.3, we give a high-level overview of our contributions while detailing the outline of this dissertation.

## 1.1 Motivation: Evolving the Industrial Landscape

Digitization and digitalization contribute to the ongoing success of the Internet of Things (IoT) [SJK+19] with its widely-distributed computing and sensing capabilities [AIM10]. In addition to the IoT consumer segment [AIM10], e.g., smart homes and their automation, digital health, or its use in transportation systems, corresponding changes also affect businesses as part of the IIoT [SSH+18]. Now, stakeholders are able to monitor production processes [SWW15, SHH+21], logistics operations [TDKCK22], agriculture and the excavation of natural resources [LKG+18, MSS+21], as well as the usage cycle of products [SSH+18]. As a result, we observe significant changes to the industrial landscape: Nowadays, companies can ubiquitously sense, process, and store an immense amount of data, covering production processes and products alike. Given the interdependencies and business relationships in industry, these developments significantly influence the cooperation of stakeholders.

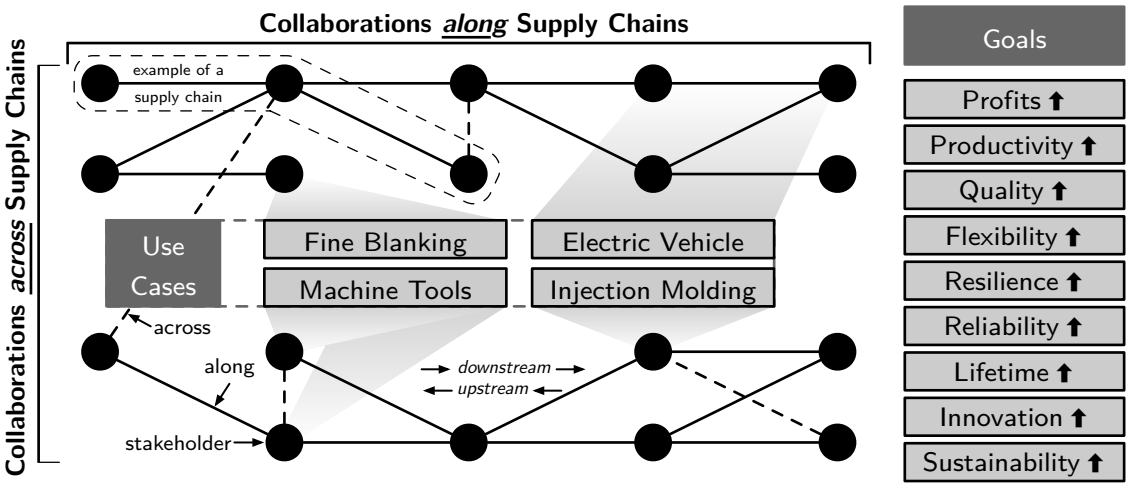



**Figure 1.1** We differentiate between collaborations of stakeholders along and across supply chains. We underline the practical relevance of our work by referring to several representative use cases throughout this dissertation. Overall, collaborations with their exchanged information allow stakeholders to better optimize their processes toward their respective operational goals.

The joint utilization of said data, i.e., *information*, therefore allows for (global) improvements of products, production processes, and manufacturing schedules, among others [SSH+18, Mou22]. Consequently, the implications of adjusting organizational and operational processes are not constrained to a single stakeholder anymore, i.e., they likely influence full product lifecycles and their associated supply chains. Especially change requests and small-batch production increasingly call for dynamic and short-lived business relationships, impacting several stakeholders at once.

Accordingly, we focus on the implications of *information sharing* and *information utilization* among stakeholders in the industrial landscape. We refer to those practices as collaborations. As we detail in Figure 1.1, in this dissertation, we look at both collaborations along and across supply chains while considering several representative real-world use cases, such as fine-blanking lines [GHW+19], the assembly of electric vehicles [KKWF16], machine tool shopfloors [Bre12], or mass production featuring injection molding [LH21]. In addition to the traditional goals of businesses like an increase in profits, productivity, and product quality [Gil16], stakeholders from the industrial landscape are increasingly interested in less tangible goals [Gil16, SSH+18, LMS+21, Mou22], such as improved flexibility, resilience, reliability, product lifetimes, innovation potential, or sustainability of their operations, as we summarize in Figure 1.1. All of them have in common that a (global) utilization of information promises to identify potentials, which, in turn, allow for process adjustments that impact those goals [BOAA+22]. Thus, the expansion and shift of these goals and associated priorities mandate the exchange of information to utilize globally-available information to the fullest extent, i.e., the use of collaborations.

### Collaborations and the Exchange of Information in the IIoT

With the IIoT-induced evolution of the industrial landscape in mind, we now elaborate on today's practices of information sharing and collaborations in industry.

When looking at today's information exchanges, the main focus is on collaborations along the supply chain [RFJ18]. Companies increasingly source any available data to extract meaningful information that influences their decision-making and operations. As a result, stakeholders try to optimize their processes, which also includes realizing short-term benefits, e.g., by being able to source goods from a different set of suppliers—for example, to lower their purchase expenses. However, recent disruptions, such as climate change [Les23], the COVID-19 pandemic [SAE23], the Suez Canal obstruction following the grounding of Ever Given in 2021 [LK23], or the Russian invasion of Ukraine in 2022 [WS23], have had a lasting impact on the production and distribution of goods and materials. Accordingly, we observe an evolution toward dynamic business relationships to swiftly react to new circumstances, i.e., to improve the companies' resilience against disruptions and other challenges in industry [SSS23]. However, these relationships are not limited to supply chains of specific products, i.e., vertical collaborations [Bar04]—*along*—the supply chain. Instead, due to the widespread yet local availability of relevant information and the manageability of sharing and exchanging said information, horizontal collaborations [Bar04]—*across*—supply chains begin to emerge. In contrast to the primarily product-focused view along supply chains, exchanging information across supply chains usually concerns more process-oriented aspects, such as the operation or commissioning of production lines (where the involved stakeholders each have their own set of experience). Given today's lack of information sharing and collaborations across supply chains, the corresponding (and currently-underutilized) information offers room for significant improvements. Thus, the development and establishment of collaborations across supply chains increase the overall importance of *coopetition*, i.e., competing businesses that cooperate for individual advantages [BDRW19, LRWW20, PTM+21].

Overall, we observe a gradual transition from locally-isolated data and information silos, where data and information are possibly not even extracted from, to globally-available *knowledge* that allows companies to benefit from the experience and findings of other stakeholders [UMNE+16]. In other words, companies begin to collaboratively adapt their (individual) operations by rigorously exchanging data and information [TBP23]. Here, a significant challenge is how to provide stakeholders with a secure, technical foundation for establishing and continuously utilizing collaborations in the IIoT that addresses both types of industrial collaborations, i.e., *along* and *across*, alike. Legislation further fuels these developments as it requires the landscape to change: While the need to document products and to standardize production processes is already a common and long-lasting practice in certain domains, e.g., manufacturing in the aerospace industry [Gor00], other domains are only beginning to digitally capture their operations [JHC21]. In light of sustainability activities and the striving for fair-trade products, legislation increasingly regulates the extent of information-producing companies need to provision. Here, Germany's 2021 act on corporate due diligence obligations in supply chains is a prime example [Bun21] as it requires companies to ensure compliance with human rights *along* their entire supply chain. Thus, the corresponding documentation from various stakeholders needs to be gathered or shared, processed, and forwarded to adhere to this legislation.

Since collaborations promise significant benefits (cf. Figure 1.1), which eventually also boil down to monetary or societal value, various initiatives evolve around them.

### Ongoing Data-Sharing Initiatives for the IIoT

We identify several data-sharing initiatives that could have a (lasting) impact on the technical foundations of industrial collaborations and, as such, the evolution of the industrial landscape. These abstract, large-scale initiatives, e.g., IDS [OAC+16, OJ19], GAIA-X [BFRLG21] or ALICE [PTM+21], promise to develop and offer new (data) ecosystems. Their aim is to support the aforementioned transition and to provide companies with technical infrastructures to rely on. Unfortunately, these initiatives are still mostly on a conceptual level, i.e., they fail to convincingly demonstrate their feasibility on a wider scale with a multitude of different use cases across domains (focusing on the information instead). Still, they relate to this dissertation from a methodological perspective. In particular, we identify several open questions that (a) arguably have relevance for industrial collaborations in the IIoT while (b) exceeding the scope of this dissertation (with its focus on information security and technical guarantees). As an intuition, the following enumeration captures the most important questions: 1. Are their collected requirements accurate? 2. Are their proposed suggestions realistic? 3. What are the incentives for companies to implement the corresponding concepts into their operations? We consider these open questions when discussing the contributions of this dissertation. However, we refrain from answering them in regard to the mentioned conceptual data-sharing initiatives as we focus on the *technical foundations* of industrial collaborations instead.

The motivation to implement and secure collaborations in industry originates from the vision to eventually come up with more sophisticated applications and use cases that greatly impact the mentioned goals (cf. Figure 1.1), i.e., allowing stakeholders to benefit from vast amounts of (newly available—given increasingly-implemented collaborations) information. Initially, today's technical building blocks from computer science provide the required security guarantees. When targeting more advanced applications, the success of future developments then also depends on goal-oriented technical advances in the domain of computer science, particularly-suitable approaches that fully address all security and privacy concerns stakeholders might have. Simultaneously, these technical advances might also spark ideas for novel, currently-inconceivable applications or use cases in the future, i.e., we expect bi-directional benefits. At this point, we want to highlight the concepts of collaborative manufacturing [McC02], digital factories [CMP+09], outsourced manufacturing [Mom02], plug-and-play parts or standardization [PNT99], and companies complementing each other [BHS05]. Moreover, the expected evolution not only affects existing businesses. With suitable technical building blocks, new business models are likely to emerge [EW17]. While the specifics are out of scope for this dissertation, a corresponding, straightforward example is when third parties offer sophisticated Supply-Chain-Management-as-a-Service (SCMaaS) practices to improve supply chain operations and processes simply as a service for others [BV19].

### Use Cases from Production Technology and Applications Beyond the IIoT

In comparison to past and ongoing initiatives, we consider the crucial aspects of suitability and technical readiness for real-world deployments. In this dissertation, we

ensure the real-world applicability of our work (in terms of performance, reliability, scalability, security, and confidentiality) by considering representative use cases from the domain of production technology. In particular, as indicated in connection with Figure 1.1, we evaluate our contributions based on their application to fine-blanking lines, the assembly of electric vehicles, machine tool shopfloors, or mass production featuring injection molding. Thereby, we also tackle the questionable practical relevance of the discussed ongoing data-sharing initiatives. Our selected use cases serve as prime examples for the challenges of securely-realized industrial collaborations due to large data volumes in conjunction with strong security and privacy needs of the involved stakeholders. Consequently, this dissertation and its contributions also greatly align with the research vision of the interdisciplinary Cluster of Excellence "Internet of Production (IoP)" [JSB+18, PGH+19, BDJ+22]—the only government-funded research cluster in Germany at the intersection of production technology and computer science, involving more than 200 researchers from different domains—which intends to eventually turn the vision of globally collaborating, yet also competing, stakeholders into reality to advance production technology, its processes, and products overall, in part by postulating the concept of digital shadows [LJ20, LJ23, vdAJKQ23]. Thereby, our work takes part in shaping the evolution of the industrial landscape through our unique focus on secure collaborations [Pen24].

Despite our focus on use cases from production technology, our ideas and approaches translate across domains due to the specifically-demanding requirements in said field, i.e., our work is beneficial to any setting where secure collaborations are of interest, including agriculture, the excavation of natural resources, or the traditional IoT with smart homes and digital health. Importantly, the expected benefits significantly exceed beyond finance, as we showcased by highlighting the goals' diversity (cf. Figure 1.1). While finance criteria like product costs, lifetimes, and quality are arguably very important to businesses, all concepts that enable secure collaborations also support the evolution of and focus on softer factors, such as environmental, social, and corporate governance (ESG) goals [FBB15], which are also expressed within the United Nation's Sustainable Development Goals (SDGs) [Gen17]. Conceptually, the exact optimization goal of a collaboration is irrelevant for the used technical building blocks and approaches that we propose in this dissertation, as the goal is unrelated to the processing. Our contributions are general-purpose approaches, and as such, they process data and information irrespective of the underlying semantics.

With this motivation for the evolution of the industrial landscape in mind, we next give a brief overview of relevant aspects when developing and implementing collaborations and, more importantly, the associated dimensions when securing them.

## 1.2 Collaborations in the Industrial Internet of Things

In the following, we first give an overview of secure collaborations and highlight the different dimensions involved when proposing and implementing them (Section 1.2.1). Subsequently, in Section 1.2.2, we briefly discuss the corresponding security and privacy challenges for our work, i.e., aspects that still interfere with the establishment of (novel) collaborations in industry.

### 1.2.1 The Status Quo in Securing Industrial Collaborations

To date, the establishment and operation of industrial collaborations are not to be underestimated and constitute cost-entailing challenges [JHC21]. We argue that the corresponding challenges follow from a lack of appropriate and secure technical solutions. Thus, many stakeholders frequently resort to business relationships with known parties to exploit existing trust relations instead [NWL10]. Moreover, collaborations in the IIoT are frequently established only after negotiating legally-binding contracts in lengthy contractual processes [FRPO15], i.e., trust is built on an organizational level and thus ensured legally rather than technologically. Naturally, this situation negatively affects stakeholders by preventing them from swiftly and dynamically reacting to disruptions (cf. Section 1.1) as short-lived yet trusted collaborations cannot be established, failing to realize significant potentials and benefits.

So far, business relationships are mainly established to be maintained long-term as the task of implementing based-upon collaborations is a lengthy and costly process, requiring significant effort by the involved stakeholders. As a result, stakeholders usually restrict their collaborations to information sharing with direct business partners within their (static) supply chains, i.e., the potential of (indirect) collaborations over multiple hops (multi-hop collaborations, from here on) or collaborations across supply chains are rarely explored. Thus, as today's status quo, companies commonly align their collaborations directly with established flows of physical products.

Following the emergence of the IIoT, trust and any associated concerns now primarily cover sensitive digital data and information on production processes and products rather than traditional, paper-based documentation. Reservations against information sharing are still deeply rooted in stakeholders due to fears of data leaks and loss of control over sensitive information [Bit23, LL23]. Hence, this attitude also impedes the widespread establishment of industrial collaborations. This stance by potentially-competing stakeholders does not come as a surprise, given that data is frequently named as the new oil [The23, MSCD18], i.e., a crucial resource for businesses and their development. With competitive advantage nowadays increasingly manifesting itself through digital capabilities, software, machine learning models, and configuration parameters [AN18]—at the expense of specialized production equipment or large practical research facilities—this attitude and behavior appear to be reasonable from a revenue-oriented business perspective. Concepts such as additive manufacturing [HJD22] significantly contribute to this shift as the production machine is not bound to specific products anymore but primarily depends on software and the supplied digital product models. In this regard, German businesses, for example, exceptionally fear counterfeit products and copycats from China [vGPZ22]. Consequently, the establishment of new and innovative yet secure collaborations (especially with untrusted companies) is still significantly impaired in industry today. Hence, in this regard, we identify the need for elaborate research activities.

Establishing and utilizing collaborations in the IIoT involves various dimensions with different foci, which we discuss in the following. As stakeholders raise various security and privacy concerns when dealing with industrial collaborations, we identify the need for reliable information security, preferably through technical means, as a key aspect. In this dissertation, we thus focus on this crucial dimension.

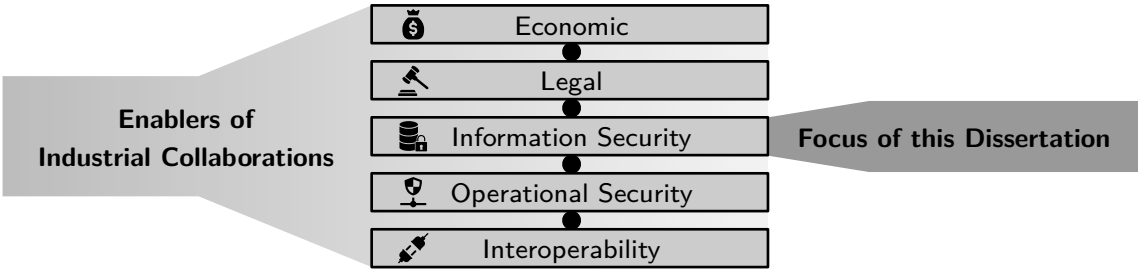



**Figure 1.2** Different dimensions impact the establishment and operation of industrial collaborations. We identify the *information security* dimension as the technological centerpiece, with significant influence on the other dimensions, and thus focus on it in this dissertation.

### The Enabling Dimensions of Industrial Collaborations

As we illustrate in Figure 1.2, we discover—at least—five mutually-influencing dimensions when realizing industrial collaborations, namely: (i) Economic, (ii) legal, (iii) information security, (iv) operational security, and (v) interoperability. We classify the information security dimension, which addresses the stakeholders' reliability and confidentiality concerns, as the technological centerpiece of industrial collaborations. Its specifics are a prerequisite when discussing the details of the other dimensions as they provide them with the technical foundation and thus greatly influence their operations. Hence, developments in the information security dimension have significant implications on the other dimensions. Consequently, in this dissertation, we focus on the information security dimension.

**Economic.** For businesses, the *economic* dimension is very important for any future developments as stakeholders generally follow monetary interests. While other criteria (e.g., ESG goals) might be relevant as well, liquidity is ultimately essential in most economies. Thus, collaborations either (i) must have a positive (monetary) impact on the business or (ii) trade this aspect off with other benefits or beliefs.

**Legal.** Likewise, the *legal* dimension is essential when implementing collaborations in practice. Stakeholders are bound to specific legislations and must be held accountable to enforce honest behavior (due to their pursuit of monetary interests; cf. economic dimension). Especially the sharing and utilization of external (third-party) information introduce novel challenges concerning liability claims [LGS17, MHS17]. In addition to physical threats to machinery, these claims are also relevant in light of safety or environmental incidents [AFS+22]. Even for legal scholars, dealing with all these challenges is virtually impossible at this point because novel collaborations and applications are neither deployed in the wild nor proposed in full. Moreover, only a few verdicts can serve as reference. Consequently, estimating the legal implications and proposing corresponding legislation is still an upcoming, long-lasting task.

**Information Security.** Furthermore, we have a dimension that covers *information security*, which concerns all aspects related to the sharing, processing, and utilization of information as part of or following collaborations. Enforcing access control, ensuring data minimalism, or protecting sensitive information are examples in this context. Consequently, this dimension primarily deals with technical challenges:

On the one hand, we require concepts to integrate collaborations into established business practices. On the other hand, we need to realize them securely. In particular, we consider the concepts of security- and privacy-by-design as well as -by-default [ISV19, SCS+21] as essential when further developing and proposing application areas of industrial collaborations. Accordingly, we classify this dimension as the technical centerpiece of collaborations in the IIoT. Future evolutions of this dimension, in accordance with the economic and legal dimensions, should address the trust concerns—that stakeholders still have these days when information flows and collaborations are involved—through technical means.

**Operational Security.** Additionally, *operational security* constitutes an orthogonal dimension of how to secure the interconnectedness of devices, sensors, and computing hardware in the IIoT [PKM18, TDDFD20]. These aspects are crucial for the automated and autonomous exchange of information within collaborations. Threats, such as denial-of-service attacks as well as any threats related to the extraction of information through compromised devices, concern involved stakeholders [PKM18]. This dimension also covers everything related to the safe operation of processes and cyber-physical systems (CPSs), i.e., safety, for the workforce and the environment alike. Especially the attack surface of legacy production devices that are nowadays (openly) connected to the Internet as well, in line with the ideas of the IIoT, is a significant issue. In this regard, research has repeatedly shown that companies operate plenty of Internet-facing devices insecurely [DLF+20, DLP+22], partially relying on insecure legacy protocols. The importance of this dimension is also apparent from the multitude of attacks in the past, e.g., an attack on a German steel plant [KNR+22] or NotPetya penetrating the operation of a large shipping company [Gre23]. However, it rather entails a secure foundation for establishing secure collaborations, i.e., it is only indirectly related to securing the collaborations themselves.

**Interoperability.** Lastly, we identify *interoperability* as a dimension: Corresponding interfaces and standards related to industrial collaborations are needed [HAAS23]. They encompass details on information flows, information processing, information systems, and associated research challenges for data management alike [JBSR17, GPL+20, GVC+22]. The mentioned large-scale initiatives (cf. Section 1.1) already tackle this dimension as part of their efforts, as their goal is to push standardized solutions and tools into the industrial landscape. We consider this dimension to primarily constitute a functional requirement, i.e., an engineering challenge, and thus, identify little need for original research within the scope of this dissertation.

This overview of the mutually-influencing enablers of industrial collaborations again confirms that the lack of appropriate solutions and concepts in the information security dimension prevents the widespread establishment and deployment of collaborations in industry so far. Next, we thus particularly focus on this dimension.

#### Information Security: The Technological Centerpiece for Collaborations

The information security dimension serves as an essential technical foundation of collaborations and is especially important to resolve the prevalent reservations and concerns against collaborations and coopetition in industry. Thus, it significantly

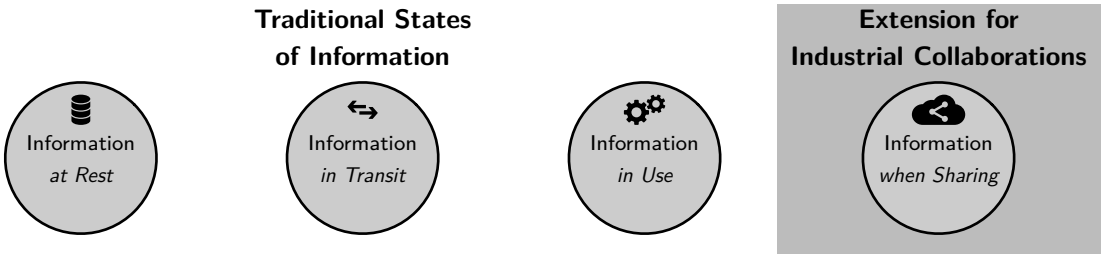



**Figure 1.3** We focus on information flows in the IIoT and define the states of information according to the traditional states of data. Given the prevalent security and privacy aspects when sharing information in collaborations, we explicitly highlight *information when sharing*.

influences the other dimensions, as outlined above. The prevalence of the IIoT, with its globally-distributed data and information, greatly contributes to its importance. Any improvements in this dimension further promise to reduce the complexity and overhead of establishing new collaborations. Accordingly, this dimension is a key aspect in the evolution of the industrial landscape. Consequently, given its importance and its implications on the other dimensions, in the remainder of this dissertation, we specifically look into information security-focused designs and contributions.

In the long run, we expect that novel, more sophisticated, and complex applications for various use cases across domains will develop following the widespread use of simpler and widely-relied-upon secure collaborations. Consequently, the first steps toward securing and improving collaborations are essential and crucial to also allow for more complex deployments. With this dissertation, we focus on exactly this task. This dissertation's goal is to serve as an important foundation to fuel and strengthen future deployments as well as (novel) applications and collaborations in industry.

### 1.2.2 Research Challenges: Security and Privacy Needs in the IIoT as Prevalent in the Information Security Dimension

In accordance with the information security dimension, we now focus on security- and privacy-related research challenges in the context of industrial collaborations. To this end, as a foundation for our work on information sharing, we first introduce the different states of information. Subsequently, we introduce the overarching research question of this dissertation, which emphasizes secure collaborations in the IIoT.

The three states of data are commonly known as *at rest*, *in transit*, and *in use* [KZ17]. In the context of this work, we generally focus on information flows as part of industrial collaborations. Accordingly, for this dissertation, we analogously define three states of information, which translate to the three states of data, as we illustrate in Figure 1.3. Moreover, as inspired by the large-scale data-sharing initiatives (cf. Section 1.1), we also consider data sovereignty [PS17, Jar20, HBTD21] in the context of this work. More specifically, we also transfer the idea of data sovereignty to information sovereignty to express it in the context of information flows rather than for individual data and datasets. Consequently, we extend the three states of information with a surrounding aspect called *when sharing*. In line with data sovereignty,

this aspect expresses the needs of shared information in the context of (industrial) collaborations, i.e., supplementary requirements related to sovereignty, which have to be considered when designing and implementing information flows.

To secure collaborations throughout the complete flow of information, i.e., across all states of information, we need to deal with the corresponding security and privacy needs. That is, stakeholders should be (i) fine with relying on secure collaborations, (ii) aware of the provided security and privacy (confidentiality) guarantees, and (iii) possibly even have a desire to advocate for secure collaborations. As we introduced before (cf. Section 1.2.1), these security and privacy needs mostly stem from trust issues and concerns with respect to today's means for information sharing.

The particular needs are highly dependent on the involved stakeholders[1], the use case at hand[2], the processed information, and the benefits a collaboration promises. An example could be to share information on the product only with actors along the supply chains, i.e., the corresponding information should not be accessible by third parties (e.g., if a cloud is involved in the collaboration) or competitors. Following the presentation of essential background information in Chapter 2, in Chapter 3, we will introduce our representative use cases along with their overall security and privacy needs to provide a more profound intuition. In our contributions (Chapters 4 and 5), we then deal with these needs in use case-specific ways while referring to the general implications and takeaways of our work for secure collaborations.

### Primary Research Question: Enabling Industrial Collaborations

With this dissertation, we intend to contribute to paving the way for secure and reliable information sharing in the IIoT by establishing secure collaborations. Specifically, we tackle aspects that large-scale orthogonal initiatives have mostly overlooked so far (cf. Section 1.1). Accordingly, our overarching research question is as follows:

*How can we enable secure industrial collaborations in real-world settings?*

Given the broad and complex nature of industrial collaborations that affect several dimensions (cf. Figure 1.2) and diverse use cases, we keep the following information security-related subquestions in mind when discussing our contributions:

- Which technical means should we source for collaborations along supply chains?
- Which collaborations across supply chains can we realize securely using today's technical means while also convincing stakeholders to participate?
- Which (technical) aspects hinder collaborations across supply chains so far?

We address these questions by proposing new *reliably-secured* designs for all relevant types of collaborations in the IIoT. We demonstrate the realistic *real-world impact* of our contributions using representative use cases from the domain of production technology. Next, we briefly present these contributions as part of our outline.

[1] In Chapter 2, we dissect the relationships of actors in the industrial landscape.
[2] In Chapter 3, we elaborate on our use cases (cf. Figure 1.1).

## 1.3 Dissertation Outline and Contributions

Within the scope of answering the overarching research question, the goal of this dissertation is to show and realize different applications as a contribution to improve the (technology) acceptance of technically-secured collaborations, exemplified in the domain of production technology. Specifically, we conduct use-inspired basic research [Sto97] to combine the fundamental understanding of secure industrial collaborations with their practical applicability. Thereby, we support and enhance the exchange and sharing of industrial information as part of secure collaborations while addressing real-world use cases. Accordingly, our work contributes to the discussed evolution of the industrial landscape (cf. Section 1.1), as exemplified by the IoP.

In this dissertation, we sort our four primary contributions by the increasing potential of the expected benefits for the overall industrial landscape once the respective approaches are widely deployed in the wild. We illustrate this structure in Figure 1.4 and label our contributions with ①–④. In contrast, the reservation against more substantial collaborations increases as well, primarily due to concerns about sharing and processing sensitive business information. From a technical point of view, securely realizing these approaches becomes more difficult as well. Overall, as we also illustrate in Figure 1.4, we group the presentation of our work into two overarching categories, collaborations along supply chains (vertical collaborations) and collaborations across supply chains (horizontal collaborations). Within this framework, we individually look into one setting with known parties and one setting where collaboration partners are unknown upfront. Thereby, we cover a large range of different scenarios with greatly-varying challenges. While collaborations along supply chains are partly established (with the potential for extensions and improvements), collaborations across supply chains promise the most significant and untapped benefits.

In the following, we first highlight our four primary contributions (①–④) in more detail. Subsequently, we briefly point out the importance of a research methodology that is suitable and appropriate for interdisciplinary research collaborations. Finally, we express the outline of this dissertation (cf. Figure 1.4) in written form.

### Collaborations Along Supply Chains

In this category, building on technical building blocks, we make two contributions to answer our primary research question: one focusing on *information sharing along supply chains* and one *tackling the challenge of identifying suitable suppliers*.

First (①), as we present in Chapter 4.1, we design an information-processing pipeline that introduces end-to-end-secured sensing to the industrial landscape. While we utilize special sensors to allow companies to guarantee the authenticity and correctness of sensed information through technical means, our blockchain-backed storage design ensures the long-term availability and verifiability of this information. Moreover, our design also efficiently supports arbitrary information sharing over multiple hops, i.e., it also covers indirect information sharing. Simultaneously, we also account for the confidentiality needs of the involved companies. As such, we provide a scalable and

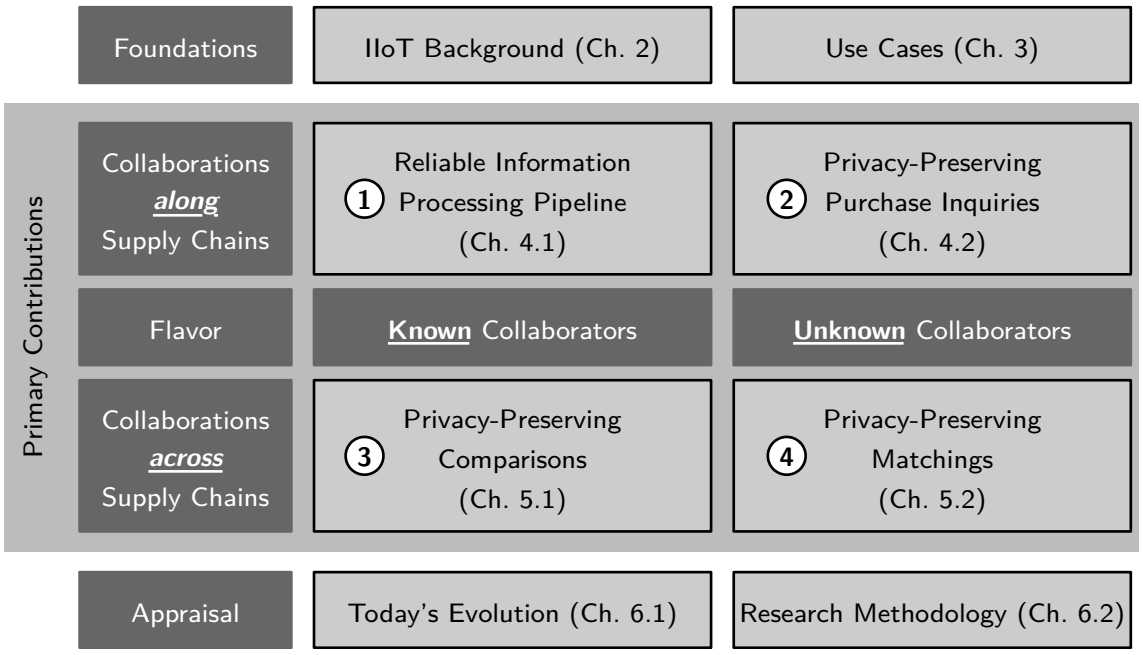



**Figure 1.4** This dissertation is structured as illustrated. We order our primary contributions by an increasing potential of benefits for the industrial landscape (①–④). Simultaneously, from an information security dimension, the prevalence of security and privacy challenges rises, and with it, the difficulty in realizing the respective collaborations (and thus, our contributions).

flexible architecture for secure collaborations along established supply chains and their stakeholders. Thereby, we enable diverse collaborations along supply chains.

Second (②), as we detail in Chapter 4.2, we address the challenge of finding new suppliers without the need to share or disclose any sensitive information upfront, i.e., before engaging in the actual collaboration, by proposing two designs with differing privacy guarantees that can be selected according to the confidentiality needs of the domain and involved stakeholders. Concerning our research question, this contribution is essential to support the establishment of additional (new) business relationships and, in turn, secure collaborations along supply chains through technical means. To the best of our knowledge, we are the first researchers to tackle this problem. Thereby, we allow companies to dynamically and privacy-preservingly identify suitable suppliers. The importance of this ability increases with the evolution of the industrial landscape as it promises to improve the overall business performance and companies' resilience. Our novel designs seamlessly integrate into today's well-established procurement processes. Consequently, we greatly support the interaction of companies with potential business partners along the supply chain.

### Collaborations Across Supply Chains

In the second category, we also make two important contributions. Overall, collaborations in this category are primarily hindered by trust issues of the involved stakeholders as they fear for their competitive advantage when sharing (or exchanging) sensitive information with competitors. To address these concerns, we propose designs that reliably ensure confidentiality in these settings through technical means.

We are confident that with the emergence of secure and trustworthy designs, the benefits outweigh the risks of unintentionally sharing (or leaking) sensitive information, even for very cautious stakeholders. Collaborations across supply chains generally differ from collaborations along supply chains as (i) trust relationships and information sharing are not yet (widely) established and (ii) they require a change in the stakeholders' mindsets, i.e., stakeholders need to grasp the benefits of coopetition rather than seeing the risks that are associated with shared information.

While our third contribution is exemplary for privacy-preserving comparisons among stakeholders, as our fourth contribution, we propose a platform that enables privacy-preserving matchings of sensitive information in industry. Comparisons as part of industrial collaborations cover applications that result in insights without feeding external information directly into local processes, i.e., such collaborations across supply chains are the least invasive as they have no direct implications for established processes. In contrast, stakeholders participate in matchings to retrieve information that can be directly implemented in their current processes. For example, configurations related to the most productive handling of identical machines might be of interest to manufacturers (possibly even competitors) across supply chains.

First (③), as we discuss in Chapter 5.1, we propose designs for company benchmarking that consider the privacy needs of both participants and the operator of the benchmark. Thereby, we close a gap in previous work that primarily focused on only one of these needs as the confidentiality concerns of the benchmarking operator were usually neglected. With our work, we correct this misconception and provide companies with a tool to identify shortcomings with respect to business partners and competitors. Here, our primary focus is on comparing the performance to competitors and companies in the same domain. With this contribution, we demonstrate the suitability of established technical building blocks to also secure collaborations across supply chains. Hence, to answer our research question, we consider respective approaches adequate to enable even challenging collaborations in real-world settings.

Second (④), as we highlight in Chapter 5.2, we realize a completely new use case that allows for privacy-preserving matchings of arbitrary information across different stakeholders in industry. While practitioners envisioned the exchange of production process parameters before [JBSR17], the lack of oblivious and privacy-preserving platforms prevented such a desired application in practice so far and, thereby, hindered the establishment of corresponding collaborations. Stakeholders expect that the global utilization of knowledge positively affects various goals, such as profits, productivity, product lifetimes, sustainability, and more [PHS+19, PMK+21]. The need for a secure and privacy-preserving exchange of sensitive information across supply chains mandates the use of technical approaches to reliably protect respective collaborations. In this dissertation, we thus propose different variants of a general design with varying privacy guarantees and performance to securely enable such desired collaborations. Our work is independent of specific information, i.e., it is widely applicable, even across domains, demonstrating how to reliably enable even sophisticated secure collaborations across supply chains in real-world settings. Thus, we account for use case-specific needs and propose an important tool to share and exchange all kinds of information with previously-unknown business partners.

### Interplay of Contributions and our Research Question

With our contributions, we present secure collaborations for all relevant scenarios in the IIoT (along and across supply chains as well as with known and unknown collaborators). Based on our work and our evaluations of practical use cases, we conclude in Chapter 6.1 that established technical building blocks are able to reliably secure collaborations in real-world settings while scaling to industry needs. Hence, we provide a technological blueprint for secure and reliable collaborations, which will emerge considerably as part of the upcoming evolution of the industrial landscape.

### A Research Methodology for Interdisciplinary Advances

We further raise the issue that developments are also needed on a methodological level: A goal-oriented and progressive evolution of collaborations in the IIoT can only be realized through deeply-rooted interdisciplinary research, even in the presence of an appropriate technological blueprint. Accordingly, a better understanding and formalization of such interdisciplinary efforts are likely to improve the achieved results, primarily due to a structured allocation of tasks and duties. We address this matter by presenting an abstract research methodology for interdisciplinary cooperations in Chapter 6.2. We derive this process cycle based on our experience that largely relate to the contributions that we present in this dissertation.

### Dissertation Outline

Before we detail our primary contributions in Chapters 4 to 5, in Chapter 2, we first provide relevant background information on the actors in the IIoT, the relevant properties when securing collaborations, and potential building blocks to realize collaborations securely. Subsequently, in Chapter 3, we introduce selected, representative use cases with distinct requirements from the domain of production technology to later evaluate our contributions with. Following our four technical contributions (Chapters 4.1, 4.2, 5.1, and 5.2), which we have just outlined (①–④), our appraisal in Chapter 6 discusses the state of the described evolution so far before elaborating on the need to also conceptually revisit interdisciplinary research collaborations as part of an excursus. This excursus sources our experience while conducting the work for this dissertation and provides researchers as well as practitioners with an abstract yet formalized process cycle to tackle interdisciplinary research challenges. Finally, in Chapter 7, we conclude this dissertation.

# 2

# Background

The introduction of this dissertation (Chapter 1) already highlighted the motives of stakeholders who will eventually implement and utilize industrial collaborations. In the following, we thus look at the corresponding industrial landscape in more detail. First, in Section 2.1, we define the term "Industrial Internet of Things" before giving a more formal overview of its main actors and conceivable information flows. Second, in Section 2.2, we summarize aspects that are relevant when securing collaborations in the IIoT. Afterward, in Section 2.3, we give an overview of the most important security concepts for our contributions, namely, privacy-preserving computation, confidential computing, and blockchain technology. Finally, in Section 2.4, we survey potential building blocks that are beneficial when addressing the relevant properties.

With this chapter, we provide the foundation for the remainder of this dissertation, i.e., (i) to showcase representative use cases for our work (Chapter 3) and (ii) to introduce our contributions using fixed terms and taxonomies (Chapters 4 to 5).

## 2.1 The Industrial Internet of Things

In this dissertation, we center our research and contributions around information flows in the industrial landscape. As we already pointed out in Section 1.1, we can distinguish collaborations along and across supply chains. More broadly, the term Industrial Internet of Things (IIoT) refers to all sorts of developments in a networked industrial landscape. We discuss the multitude of definitions of this term in Section 2.1.1 and put them into perspective with the focus of this dissertation. Subsequently, in Section 2.1.2, we give an overview of the different actors in the IIoT before discussing different topologies that are suitable when establishing industrial collaborations in Section 2.1.3. This background creates the foundation for (i) our discussions on their relevant security properties and (ii) our survey on building blocks that come to mind when securing industrial collaborations in real-world deployments.

### 2.1.1 Definitions and Taxonomy

On a general level, the IIoT covers a wide range of applications, technical concepts, and future roadmaps under the umbrella of networked industrial devices. As a result, various definitions emerged [BHCW18]. We refer to Boyes et al. [BHCW18] for a detailed comparison of different viewpoints in research. In the context of this dissertation, we follow their conclusion. That is, we consider the IIoT to cover all facets related to the industrial utilization of the IoT, CPSs, networked devices, and information technology in general. Consequently, the IIoT captures a paradigm change as, traditionally, devices and machines have been (locally) isolated, i.e., their networking was severely constrained (and underutilized). Hence, it increasingly contributes to the convergence of operational technology and information technology [BHCW18]. Just as in the IoT, the IIoT intends to benefit from a large number of sensors that generate vast amounts of information for real-time computations, post-processing, and long-term use alike.

Various work [BHCW18, QCZ+20, KNT+21, PBB+23, RAB+23, SKTG23] looks into the appropriate placement and operation of networking and computing infrastructure to enable the increasing utilization of information. In light of the IIoT and its effect on the evolution of the industrial landscape, we expect a shift from device-to-device (or machine-to-machine) communication to an interconnected setting with in-depth yet secure and reliable production-to-production communication [PHS+19]. The latter refers to production environments (which also feature CPSs) directly communicating with each other to source and process global knowledge without requiring any human interaction. Thus, such environments are destined to autonomously optimize their individual operations on a global scale, i.e., knowledge and information will be sourced, exchanged, and processed across stakeholders in real time.

Hereby, the application of the IIoT is not limited to a single phase of a product's lifecycle. Instead, it concerns all aspects of value chains in industrial settings. Accordingly, the IIoT is not confined to control loops in CPSs but also considers processing loops related to strategic or operational decisions [PHS+19]. Especially in settings with industrial collaborations, these loops also incorporate information from other stakeholders, such as suppliers, consumers, or even competitors. Accordingly, in the next subsection, we present an overview of these entities and their relations.

#### Internet of Production (IoP)

Embedded in the general context of the IIoT, our work originates from the Cluster of Excellence "Internet of Production" (IoP) as part of Germany's excellence strategy of the federal and state governments. In the IoP, academia researches the impact and implementations of digitization, IoT, and CPS for the production industry in an interdisciplinary setting. This research covers the entire product lifecycle, i.e., production, development, and usage [PGH+19, Kor20]. It is the only Cluster of Excellence (out of 57 in Germany) that conducts research in the area of production technology and the Industrial Internet of Things. Altogether, more than 35 institutes and 200 researchers are involved in this interdisciplinary research cluster.

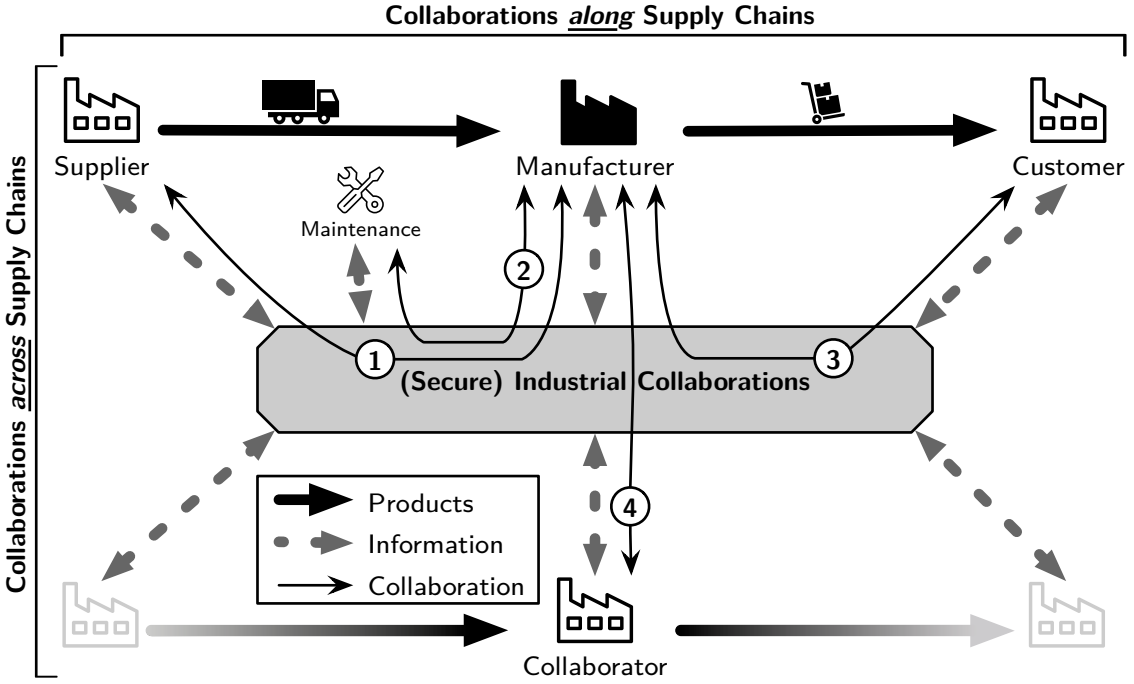



**Figure 2.1** We identify five conceptual entities in the industrial landscape that utilize industrial collaborations to exchange information. We label the respective collaborations from ① to ④.

The IoP pursues the vision to enable a new level of cross-domain collaboration by providing semantically-adequate and context-aware data from production, development, and usage in real time on an appropriate level of granularity [PGH+19]. Real-time, secure information availability of all relevant data at any time and at any place is a core aspect of this vision [BDJ+22], which calls for research on (secure) industrial collaborations and their corresponding information flows. This way, the IoP paves the way for a new era of fast, versatile, and dependable production [PMK+21].

Recent advances in the IIoT enable the measuring and extracting of massive amounts of data related to products and their production processes [GHW+19]. Sourcing the corresponding information embedded in the data and combining information from the entire product's lifecycle promises to generate new insights and streams of revenue [SKPP23]. Even more, with information automatically shared in real time by all involved entities, companies are able to minimize production interruptions independent of the responsible source, improving their local operations [PHS+19]. Currently, companies mostly rely solely on local information, effectively sealing knowledge in stakeholder-specific information silos that are not accessible by external parties at all, rendering automated collaborations infeasible [PHS+19,GPL+20]. As a prerequisite for discussing IIoT-related advances and sophisticated industrial collaborations, we continue with a presentation of the industrial landscape and its actors.

### 2.1.2 An Overview of the Industrial Landscape

In the context of collaborations in the IIoT, we consider every stakeholder to be a collaborator that can collaborate with any other stakeholder (collaborator). However,

since their relationships differ in practice, we categorize the different stakeholders into five conceptual entities [PHS+19]: *supplier*, *manufacturer*, *collaborator*, *customer*, and *maintenance provider*. This categorization allows us to better structure their relationships from the viewpoint of a specific manufacturer. Focusing on a different stakeholder changes the mapping of conceptual entities accordingly.

In Figure 2.1, we illustrate their embedding within the industrial landscape from the point of view of a single manufacturer (top center). The upper product flow refers to regular supply chains, where the manufacturer receives goods from a supplier and manufactures a new product, component, or part. This company's output is then delivered to a customer, which can be either a merchant, an end customer, or another manufacturer, i.e., the current company is a supplier with regard to the following entity in the supply chain. Any information exchange among these entities constitutes industrial collaborations *along* the supply chain (① and ③). Furthermore, we identify a maintenance provider which closely interacts with other entities (e.g., suppliers, manufacturers, or customers) without being directly involved in a product flow. We label its information flow with the manufacturer as ②. For simplicity, we only include a single maintenance provider in Figure 2.1.

Moving toward collaborations *across* supply chains, we exemplarily include a collaborator of the manufacturer (bottom center in Figure 2.1), though, in practice, manufacturers can have multiple collaborators at the same time. We label the associated collaboration and flow of information as ④. For example, two companies that operate similar production lines can exchange insights on the production performance or best-performing configuration parameters. Irrespective of the manufacturer's supply chain (and point of view), this entity is also part of its own supply chain and manages its own product flows. With the increasing emergence of (secure) industrial collaborations (as envisioned by the IoP), each entity in the industrial landscape will likely have several collaborators to exchange information and knowledge with.

#### 2.1.2.1 Entities in the Industrial Landscape

For future reference, we provide a short definition for each of the five entities in the industrial landscape. Their main properties are as follows.

- The **supplier** delivers materials, parts, goods, or intermediate products to the manufacturer (its customer along the supply chain).
- The **manufacturer** is our point of view in Figure 2.1. It is part of a longer supply chain and performs a manufacturing step using the received goods before shipping the deliverable(s) to its customer(s).
- The **customer** receives (intermediate) products, parts, or goods from the manufacturer (its supplier backward along the supply chain). When receiving a final product, the customer can also be an end customer (product user).
- The **maintenance provider** directly interacts with other entities, i.e., its clients, to perform maintenance-related tasks at their respective production sites.
- The **collaborator** acts as an entity that is contacted to benefit from industrial collaborations across supply chains. Concerning the collaboration, its operation and flow of products are unrelated to the supply chain of the manufacturer.

Given our selected point of view, we simplify real-world supply chains that consist of several entities. Thus, three of the introduced entities are also manufacturers when taking a different viewpoint, i.e., the supplier, the collaborator, and, depending on the role of the entity, the customer as well. Moreover, the industrial landscape can also host multi-hop collaborations among indirect business partners along the supply chain. While we refrain from integrating such collaborations in Figure 2.1, a collaboration (without the direct involvement of the manufacturer!) between the illustrated supplier and customer would constitute such a multi-hop collaboration.

#### Considered Attacker Model: Malicious-but-Cautions Entities

When investigating industrial collaborations and proposing new designs to implement them securely, we also have to study relevant security threats and risks. Just like in collaborations along supply chains, entities in collaborations across supply chains (possibly even among competitors in their domain) have an incentive to extract as much sensitive information as possible for their individual gain. However, given that entities and other commercial entities (e.g., third parties) in the industrial landscape are registered businesses that operate under specific legal jurisdictions, we consider them to be very cautious as misbehavior could be easily punished by law, e.g., incur huge monetary fees. Moreover, they are usually well-known in their domain. Thus, they also depend on their reputation to attract further business and generate revenue. Furthermore, whenever companies invest funds and resources into industrial collaborations, they most likely have little incentive to misbehave, e.g., by providing incorrect or wrong information to the collaboration, especially if such actions are detectable or if they diminish the usefulness of the collaboration.

Correspondingly, from a security perspective, we expect that industrial collaborations must be secure in the context of malicious-but-cautious entities [Rya14], i.e., these entities want to extract as much information as possible without leaving any traces of the extraction (and their misbehavior). Otherwise, they behave according to agreed protocols. This behavior also covers that they do not abort industrial collaborations randomly or halfway. Furthermore, given the aforementioned reasons, we consider collusion attacks that involve multiple companies as highly unlikely in the industrial landscape. Collusion attacks are attacks in which two or more parties interact to extract sensitive information they could not retrieve on their own. As most jurisdictions have laws in place that ban or strictly regulate cartels [OECD13b,HS14], we consider the likelihood of companies colluding with a third party than among each other as more threatening. Therefore, we place particular emphasis on collusion attacks that involve third parties (i.e., "introduced" entities that are not necessarily or directly linked to the industrial landscape) but still consider all kinds of collusion attacks as part of our security discussions to provide a holistic view in our work.

### 2.1.2.2 Industrial Collaborations and their Information Flows

Moving on from the conceptual entities, we now take a look at the different types of industrial collaborations in the IIoT and their respective information flows. Given

that the information covered in collaborations as well as the corresponding confidentiality needs vary significantly in practice, we discuss the different types of collaborations individually. We first consider the traditional relationships of the manufacturer along the supply chain before referring to collaborations across supply chains.

#### Type ①: Supplier and Manufacturer

Traditionally, the supplier delivers raw materials or intermediate products to the manufacturer in a unidirectional flow of products and information. One aspect of deepened collaborations among these entities is the utilization of received information. For example, the manufacturer could efficiently adapt its processes based on detailed material properties from the supplier. Likewise, suppliers and manufacturers could collaborate more closely in the development of new products to utilize synergy effects based on insights from previous usage information. However, corresponding collaborations mandate accountability guarantees because the involved entities fear improper operational and strategic adjustments otherwise. Overall, deeper collaborations among suppliers and manufacturers help to reduce inaccuracies in products (i.e., improve the quality) by adapting processes for specific workpieces, optimize the product development by combining information of entire product lifecycles, and cut down costs due to better manageable production planing and control (PPC).

#### Type ②: Maintenance Provider and Manufacturer

Collaborations among the maintenance provider and its clients (here: manufacturer) promise to improve the maintenance provider's response time. In a collaborative scenario, the maintenance provider should be able to minimize the clients' downtimes by conducting predictive maintenance and shipping replacement parts on time solely based on insight into the running processes at the client. Besides, based on the usage information, the maintenance provider can predict usage estimates or offer remote repairs that further help both entities to improve their PPC by scheduling production and maintenance times accordingly. Overall, both entities improve their product knowledge, which improves their efficiency and resilience in unexpected situations.

#### Type ③: Manufacturer and (End) Customer

Nowadays, collaborations among end customers and "manufacturers" are already well-established in the digital world [Spi12]. For example, tracking and usage information allows companies that offer (digital) services to draw meaningful conclusions. In an evolved industrial landscape, these collaboration-based advances will likely follow in the IIoT as well. On the one hand, customers are interested in receiving the best-suitable user-tailored products. On the other hand, manufacturers want to minimize unnecessary expenses by only offering and supporting needed features based on customer usage. These requirements call for industrial collaborations with agile processes as well as flexible product development.

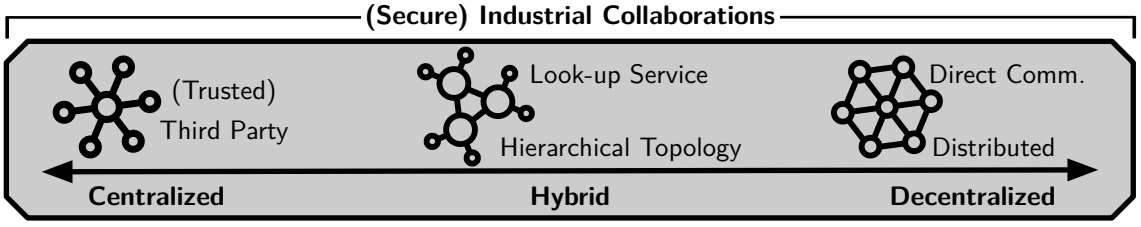



**Figure 2.2** When facilitating the organization and operation of industrial collaborations with a heterogeneous set of stakeholders with individual security needs, different concepts are realistic. Given the scale of the industrial landscape, mixing these concepts as seen fit is also possible.

#### Type ④: Manufacturer and Collaborator

Today, manufacturers mainly obtain improvements through computational advances or collaborations along the supply chain, i.e., industrial landscape mostly lacks collaborations across supply chains. However, the benefits of such collaborations are particularly interesting due to their vast amount of available experience, knowledge, and information. For example, both collaborators (manufacturer and collaborator) can simply process the same raw material or operate a machine by the same manufacturer. Sharing information can help to realize untapped potentials. In general, any entity in the industrial landscape can be a collaborator. The participating entities are not linked by a particular product or its supply chain. Given that the creation and secure realization of such collaborations across supply chains still remain unclear (among other aspects, established trust relationships are usually missing), companies rarely exploit them so far. In this dissertation, we look into these research challenges (cf. Section 1.2.2) and propose new designs to overcome them (Chapter 5).

Before looking at properties that are relevant when securing collaborations (Section 2.2), we first discuss the scope of possible real-world collaboration topologies.

### 2.1.3 Facilitating Collaborations in the Industrial Landscape

To implement the discussed industrial collaborations, different topologies and modes of operations come to mind, e.g., to account for the security needs of stakeholders or the specific circumstances of different use cases. Overall, we identify three conceptual topologies to facilitate collaborations in the industrial landscape, as we detail in Figure 2.2: *centralized*, *decentralized*, and *hybrid*. The choice can differ between different collaborations, and consequently, a single entity can be involved in different collaboration topologies. The topology can even vary depending on the information that is being exchanged. While collaborations along supply chains and with known collaborators are simpler to establish due to the existing business and trust relationship, collaborations across supply chains and/or with previously-unaffiliated entities are likely to bring significant benefits concerning the outlined goals from Figure 1.1.

**Centralized.** Such operations mandate the use of a *(trusted) third party* to realize collaborations. Given the central point of contact, such a topology eases the bootstrapping of new collaborations, and entities can be easily matched and put into contact,

i.e., no costly gossiping is needed [KVS07]. If collaborators pay or compensate the third party, it also has an incentive to operate it. However, in addition to introducing a single point of failure, this topology establishes a valuable target for data theft as sensitive information, possibly including the entities involved in a collaboration, is centrally available. Such challenges are well-known from cloud computing, and various approaches, e.g., homomorphic encryption [AAUC18] or scaling out [Gro09], to address these security and scalability issues are available [TJA10, HHH+17].

**Decentralized.** On the other side of the spectrum, decentralized topologies operate without a central entity. Here, *direct communication* among collaborators and more sophisticated *distributed networks* come to mind. While direct communication offers significant security benefits and does not introduce significant management overhead, it complicates the bootstrapping of new collaborations as entities are unaware of potential collaborators and their associated benefits. Distributed networks are well-known in the context of peer-to-peer networks [SW05] and blockchain technology [Pil16]. Given this background, it offers readily available methods for joining and departing entities or a lookup of capabilities. However, depending on the concrete realization, participating entities must trust various (unknown) stakeholders.

**Hybrid.** As a middle ground between these extrema, hybrid approaches are available. *Lookup services* offer a single point of contact to ease the bootstrapping of new collaborations before requiring collaborators to set up a form of direct communication. Thus, they are a trade-off between the mentioned bootstrapping challenges and the danger of introducing a (high-value) single point of failure. Alternatively, a federated or *hierarchical topology* can combine the benefits of a lookup service and a distributed network. In the past, by relying on so-called supernodes, multiple peer-to-peer networks used this concept to improve discoverability and scalability [SW05].

The wide range of options when setting up industrial collaborations underlines the scope of this challenge. In addition to collaboration-specific needs, each topology's usefulness and exact security properties eventually also depend on the deployed realization. Thus, real-world deployments mandate a careful and extensive analysis.

## 2.2 Relevant Properties when Securing Collaborations

When looking at industrial collaborations from the information security dimension (cf. Section 1.2.2), several properties are important to secure them. To make these properties easier to grasp, we now group related security- and privacy challenges. To this end, we rely on the well-established information security concepts of confidentiality, integrity, and availability (CIA) [WM11] as well as authentication, authorization, and accountability (AAA) [WM11]. As stated (cf. Section 1.2.2), we deliberately focus on the information security dimension in the following and do not specifically consider any orthogonal properties in the operational security dimension. We further consider the associated challenges of key distribution, key agreement, and key exchange to be orthogonal research and refer to related work [NLO15, KBL18].

The relevance for each information flow differs depending on the involved entities, the transmitted information, the chosen topology, and the stakeholder's reliability

and confidentiality preferences. At this point, we want to emphasize that secure collaborations also cover the aspects of accountability, authenticity, and verifiability, as they all contribute to the reliability of information flows and subsequently-processed information. Some particularly-cautious stakeholders might have an aversion to industrial collaborations. Hence, their needs for perceived security and privacy are higher than the needs of more open-minded stakeholders. To address these concerns, we identify three higher-order categories of security and privacy challenges: *authenticity of information*, *scope of information access*, and *anonymity*.

**Authenticity of Information.** The first category covers properties related to the correctness and origin of information. In industrial collaborations, the *authenticity* of information is vital because entities utilizing said information must be sure that it is reliable. Otherwise, wrongfully-adjusted machine parameters could incur significant damages, or jobs might be scheduled without having an actual buyer at hand. Similarly, *integrity* protection is crucial to prevent any tampering with exchanged information. Apart from ensuring the authenticity of information, establishing *accountability* is another important property. For example, a machine manufacturer guarantees the validity of its usage estimates to the machine operator. Consequently, it should be accountable. Especially more complex information flows, e.g., in multi-hop collaborations or when information is processed at several hops, impose significant challenges. Likewise, end-to-end guarantees along supply chains are imaginable as well. Here, a manufacturer of brake pad carriers could attest to the end customer of a car that the delivered component endures the car's lifetime. In the IIoT, a significant challenge is to realize a link between the physical object and its digital information because attaching a physical identifier, such as a barcode, RFID tag, or black light marker, might not always be an option. Even more, to prevent counterfeit products and other wrongdoing, such identifiers need to be tamperproof.

*Auditing* and *verifiability* capabilities are of interest in industrial collaborations to allow for verifying processes and exchanging information. In parts, this goal overlaps with the accountability property as companies should be able to prove which interactions they initiated anyway. Relatedly, *immutability* and *referenceability* enable stakeholders to make sure that the transmitted information is long-term locatable and remains unaltered (e.g., for verification purposes). Hence, all information processing should be designed with these requirements in mind. These properties integrate nicely with an evolving IIoT because companies are expected to improve their processes by applying gathered (past) knowledge. Hence, they can build on the promise that information must be retrievable and available anyway.

**Scope of Information Access.** While the first category mainly deals with the trustworthiness and reliability of information, this second category comprises different properties related to the access to information. In the context of the industrial landscape, most information is valuable because it contains details about the production process, registered patents, or created intellectual properties. Consequently, all entities have a large incentive to retain their (sensitive) knowledge locally. *Confidentiality* mechanisms, reducing the number of information flows, and limiting the extent of collaboration to a minimum allows companies to participate in secure collaborations. A reduced *granularity* of information (e.g., through aggregation,

blinding, or anonymization) can help to rule out the dangers of reverse-engineering or side-channel leaks based on shared information.

In addition, this category deals with properties related to the access to information. Proper *authentication* should ensure that no information is leaked to unintended parties, i.e., requiring each entity to authenticate itself. Furthermore, *authorization* must be granted as well to obtain access to information. While these aspects are insignificant for information flows between two entities, the challenges are more demanding when taking into account that information is expected to be forwarded along the supply chain (i.e., over multiple hops). Therefore, *information control* concepts require careful evaluation. In particular, the list of authorized entities must be expandable. For example, depending on the chosen topology, (trusted) third parties might also need access to otherwise sensitive information, e.g., for the sake of running specifically-permitted intermediate processing steps. Consequently, their capabilities must be properly defined. As preventing unauthorized information forwarding (from any entity) is a technically very sophisticated task, most regulation is likely based on contracts. When dealing with misbehavior, e.g., if information has been utilized, exchanged, or forwarded without permission, the previously-mentioned auditing property supports in identifying the wrongdoer.

**Anonymity.** The third category that we identified deals with the anonymity of collaborators. While direct business partners along the supply chain know each other, multi-hop knowledge might not be required or desired. For example, companies could strongly oppose the disclosure of their network of suppliers or maintenance providers. Regardless, their actions should still be covered by *identifiability*, providing a unique reference for each action and entity to achieve accountability. In addition, *untrackability* must be taken into account to prevent that side-channel information or communication patterns can de-anonymize participating companies. Overall, companies might require anonymity mechanisms in place to support novel, secure collaborations that complement traditional flows of information. Especially parties with an aversion of sharing information might not collaborate otherwise.

After having outlined these properties, we now continue with an overview of concepts and building blocks that could fulfill them when designing secure collaborations.

## 2.3 Building Blocks for Secure Collaborations

When securing collaborations, we can source building blocks and concepts from various areas: (i) cryptography unrelated (e.g., access control), (ii) traditional cryptography (e.g., encryption), (iii) privacy-enhancing cryptography (e.g., homomorphic encryption or secure multi-party computation) [BP21], (iv) other privacy-enhancing technologies and concepts (e.g., differential privacy or federated learning) [GSU+22], (v) authenticity-enhancing technologies (e.g., confidential computing) [SMGMM22], and (vi) recent developments (e.g., blockchain technology, smart contracts, or verifiable computing). The exact choice depends on the requirements of the use case.

In the following, we first introduce the most important building blocks for our dissertation's contributions in more detail. In Section 2.4, we then look at the entirety of

these building blocks and concepts (including the ones that we did not utilize in our contributions). To secure collaborations, we primarily rely on two diametrical strains from *private computing*, namely, the software-based concept of *privacy-preserving computation* and the hardware-based concept of *confidential computing*. Moreover, we introduce *blockchain technology*, a cryptographic approach to reach a consensus among mutually-distrusting entities—a setting that closely relates to the industrial landscape with its competing and mutually-distrusting stakeholders.

### 2.3.1 Privacy-Preserving Computation

Approaches in the area of privacy-preserving computation protect sensitive information from unauthorized access while simultaneously supporting (a few) selected operations on the protected information without generally impairing the confidentiality of the processed information. Corresponding approaches have in common that they make use of encryption. Consequently, they constitute software-based security concepts. In the following, we introduce the main characteristics of four approaches, which we apply in the remainder of this dissertation, namely, oblivious transfers (OTs), private set intersections (PSIs), HE, and order-revealing encryption (ORE).

#### Oblivious Transfers (OTs)

In the most basic form, OTs allow a client to "covertly" retrieve one of two items that are stored on a server without the server knowing which of the two items has been requested by the client [EGL85, Rab05]. After concluding the OT, the receiver has only access to a single item and does not gain (additional) insights into the other item. This basic form is also known as 1-out-of-2 OT. Several additions to this basic form enable more sophisticated retrieval scenarios for practical use, such as 1-out-of-$n$ OTs or $k$-out-$n$ OTs [CT05]. Consequently, OTs are an important cryptographic building block when designing protocols for privacy-preserving computations as they enable privacy-preserving retrievals of information.

For improved performance, a few expensive base OTs can also seed a large number of less expensive OT extensions [Bea96, IKNP03]. However, to achieve the required security and privacy guarantees, i.e., hiding the contents of the data transfer, significant computational overhead and communication are introduced [NPP01]. While the trade-off between required cryptographic computations and communication overhead is adaptable, OTs remain costly today. Consequently, they are not suitable for efficiently transferring large amounts of data in practice (as prevalent in the IIoT).

#### Private Set Intersections (PSIs)

PSIs are another cryptographic building block that allow two parties to calculate the intersection of two confidential sets without revealing the included elements to the other party [DCT10, MAL23]. Depending on the concrete implementation, only one or both parties learn the content or the size of the computed intersection [DGVPdPS12]. PSIs have been realized with different underlying cryptographic

building blocks. Nowadays, many efficient designs utilize OTs for improved security [PSZ14, KKRT16, RR17]. As a result, just like OTs, PSIs also suffer from (possibly infeasible) processing and networking overhead with increasing set sizes. For better flexibility, other variants internally rely on RSA and Bloom filters without impairing the claimed security guarantees [KLS+17]. Finally, first variants also integrate homomorphic encryption to realize PSIs [CHLR18].

### Homomorphic Encryption (HE)

HE allows for calculations on encrypted data without requiring access to the underlying raw data, thus maintaining data confidentiality even when computing on untrusted hardware [AAUC18]. Conventional encryption schemes require decrypting the data before any calculations can be executed. Even though the result can be encrypted again, the entity which executes the calculations requires access to the keys of the data owner. Thus, an offloading of calculations is not possible without abandoning data privacy in the traditional way. In contrast, HE allows the execution of mathematical operations directly on encrypted data [AAUC18]. From a technical perspective, HE schemes can have varying cryptographic foundations that differ in terms of supported operations and encrypted data types (e.g., Booleans, integers, or approximated reals), each with individual overhead and constraints [AAUC18]. Moreover, the concept of hybrid homomorphic encryption [NLV11, DGH+21] combines symmetric ciphers with HE to reduce the size of the ciphertexts, i.e., it reduces the communication overhead of HE at the expense of more complex computations.

Different variants of HE feature distinct implications on usability, precision, performance, and storage, including *partially homomorphic encryption (PHE)* [RSA78, GM84, Pai99], *somewhat homomorphic encryption (SWHE)* [BGV12], and *fully homomorphic encryption (FHE)* [Gen09, VDGHV10, BV14], as we detail next.

*Partially Homomorphic Encryption (PHE).* This subtype of HE enables (repeated) computations of a single specific arithmetic operation, such as addition or multiplication, on ciphertexts while introducing moderate computational and storage overheads [AAUC18]. For example, the Paillier cryptosystem [Pai99] supports additions of ciphertexts. In contrast, ElGamal [ElG84, ElG85] enables multiplications of ciphertexts. In contrast to more advanced subtypes, PHE schemes do not support the (repeated) computation of another operation directly on ciphertexts.

*Somewhat Homomorphic Encryption (SWHE).* To address the aforementioned shortcomings, SWHE schemes [BGV12] can compute additions and multiplications on ciphertexts. When performing these operations on ciphertexts, noise is added to the resulting ciphertext, which accumulates over subsequent operations. Thus, once the noise budget has been exceeded, correct decryption of the ciphertext is no longer possible. Hence, the number of subsequent operations is limited, which is why literature also refers to those schemes as leveled fully homomorphic encryption [AAUC18].

*Fully Homomorphic Encryption (FHE).* In contrast to the previous subtypes, FHE schemes [Gen09, VDGHV10, BV14] can compute arbitrary functions securely on the ciphertexts. Their practical applicability is thus only limited by their significant

processing and storage overheads. Especially sequential multiplications introduce significant overheads and decreased accuracy [AAUC18]. FHE schemes usually support bootstraps to overcome the aforementioned limitation in the number of repeated operations (cf. SWHE) [AAUC18]. Basically, during bootstrapping, a ciphertext is re-encrypted to "reset" the noise budget, i.e., to create room for additional operations on the ciphertexts. Usually, bootstrapping operations are quite costly [CJL+20], and thus, the specified computations should be optimized to use as few as possible.

The difference between FHE and SWHE can also surface in implementations. For example, while Microsoft SEAL [CLP17,Mic18] also implements large parts of an approximating FHE scheme, called CKKS [CKKS17], it does not support its bootstrapping feature [Lai23], rendering the implementation to (only) perform like an SWHE scheme. Looking at recent developments, CONCRETE [CJL+20], a new FHE scheme that is based on the TFHE cryptosystem [CGGI20], promises to perform bootstraps more efficiently. Moreover, it introduces the concept of programmable bootstraps, which enables the computation of univariate operations while bootstrapping a ciphertext [CJL+20]. Thus, these programmable bootstraps further extend the possibilities when performing FHE-based privacy-preserving computations.

#### Order-Revealing Encryption (ORE)

As the last approach, we discuss ORE [BLR+15], which extends the concept of order-preserving encryption (OPE) [AKSX04]. Thereby, it introduces a building block for privacy-preserving computations that allows for efficient range queries and sorting on encrypted data [CLWW16, KT19]. ORE addresses some of the limitations of OPE, e.g., to mitigate inference attacks and prevent information leakage of the "encrypted" plaintext, and accordingly intends to only reveal the order of the encrypted plaintexts [BLR+15]. Thus, as intended, when applying ORE, the computing party learns nothing except for the ordering of the ciphertexts. In the context of industrial collaborations, this feature can be useful when privacy-preservingly comparing the inputs of different stakeholders.

### 2.3.2 Confidential Computing

Confidential computing covers hardware-based security concepts that secure information by isolating it from insecure or untrusted parts of the computing environment. This paradigm shields all sensitive information in a trusted part, and thus, is also called trusted computing. Corresponding approaches rely on attested trusted execution environments (TEEs) to ensure that the computing remains confidential.

#### Trusted Execution Environments (TEEs)

First of all, TEEs depend on hardware roots of trust to realize confidential computations. Conceptually, TEEs *isolate* parts of the running software (and thus, the computation) within the device [MGDC+17, SMS+22]. Reliable *attestation* mechanisms

prove the correctness of the (i) running software and (ii) performed computations (to remote stakeholders) [MGDC+17, SMS+22]. The attestation allows the opening of an authenticated and encrypted communication channel, guaranteeing the remote stakeholder to communicate verifiably and securely with specific software on the device. This feature is particularly useful when dealing with mutually-distrusting parties or when deploying the device in untrusted environments. In addition to these two basic properties, TEEs typically also provide *memory protection* and *sealing* (protected disk storage) [MGDC+17]. The isolated and trusted part of a system is also known as an *enclave* [MGDC+17]. Some TEEs even offer direct control of and channels to peripherals, e.g., sensors, from within the enclave [MGDC+17, SMS+22].

Intel SGX [MAB+13] and ARM TrustZone [PS19] are two popular and widely-known examples that are available in modern CPUs. Depending on the hardware, their capabilities and security features can vary slightly. While the former is best known from desktop and server CPUs, the latter is popular in IoT devices and smartphones. Hence, confidential computing is also suitable for use in the IIoT. Intel SGX is also available in cloud environments, e.g., using Microsoft Azure [Rus23]. Additionally, more lightweight TEEs, such as Sancus [NBM+17], increase the range of application areas for confidential computing and even allow for widespread deployments of TEE-backed IoT devices. Such cost-efficient variants base on low-level processors with low resource capabilities. For interoperability, mutual attestations among different TEE variants and vendors are also supported [SPN+23].

### 2.3.3 Blockchain Technology

Blockchain technology [Pil16, RKC24] is a distributed ledger technology that builds on a distributed and immutable append-only ledger, which utilizes cryptography to irrevocably link *blocks* to form a chain [NBF+16]. By appending blocks, altering or removing older blocks becomes computationally infeasible, i.e., all information is stored in a *tamperproof* (immutable) manner. Initially, blockchain technology has been designed to facilitate financial transactions without requiring a trusted third party [NBF+16]. However, it can also record other assets as well as additional information that is stored with the transactions [MHH+18, MPBW20, MPW20, MPW23, MHMW24, PBW+24]. Using a consensus protocol, all participants of the system agree on a consistent state of the blockchain [Pil16]. In the context of this work, having and maintaining a consistent and tamperproof view can be very beneficial, especially when dealing with mutually-distrusting entities.

Due to the reliable long-term storage of information, blockchains are well suited to address accountability and verifiability requirements as they can guarantee the *existence* of data without relying on any trusted parties. Given that all information is irrevocably persisted on the blockchain, carefully considering which information needs to be stored is important [ZXD+18, MKP+21]. Apart from such confidentiality issues and storage overheads, another scalability issue concerns the number of transactions that can be reliably processed by the consensus protocol [SSS17]. In addition to public and permissionless blockchains, such as Bitcoin [NBF+16], that allow everyone to participate and submit transactions [Pil16], other variants emerged as well:

Private and permissioned blockchains, such as Quorum [CON20], restrict access to the blockchain to specific entities or groups [Pil16]. Thereby, they can address certain scalability aspects more easily [CP22]. Alternatively, sharding [ZMR18] (splitting the blockchain into smaller partitions), sidechains [Pil16] (that are linked to the main blockchain), and only persisting *fingerprints* (cryptographic hashes) [CG20] of the actual information are strategies to improve the scalability of blockchains.

With this discussion of three conceptual directions (which we build on in our designs) in mind, we move on to a more general survey of concepts and building blocks that promise to tackle the relevant properties we highlighted before (cf. Section 2.2).

## 2.4 Building Block Survey: Securing Collaborations

For our building block survey, we expand the range of technologies we consider as potential solutions to fulfill the relevant properties outlined before (cf. Section 2.2).

Given that we identified utility-specific clusters within these building blocks, we grouped them into five larger groups that loosely target similar challenges. In particular, we categorize the building blocks as follows: (i) *data security* covers building blocks that mainly deal with access to information, (ii) *data processing* concerns technologies that aim to conceal information during computation, (iii) *proving support* deals with mechanisms to establish authenticity of information, (iv) *platform capabilities* incorporate building blocks that ensure strict rules and foster verifiability, and (v) *external measures* contain supporting concepts that facilitate establishing industrial collaborations while not primarily focusing on security properties. Next, we introduce these categories and the individual building blocks they encompass in more detail. Table 2.1 accompanies this presentation: We provide a high-level overview of the different building blocks and the relevant properties they address.

**Data Security.** We grouped building blocks with a strong focus on the access to information, i.e., providing confidentiality, into this category. Here, the most basic form to achieve confidentiality is to rely on *encryption* [BBM00]. While regular encryption has no drawbacks concerning the other challenges we defined, it lacks features to dynamically update the number of entities that are allowed to access the information without leaking the used key or still sharing the content with removed entities (even when data is updated at a later point). In an evolved industrial landscape, relationships are more dynamic and short-lived. Accordingly, corresponding approaches should account for scenarios where access needs to be granted in a flexible manner to changing entities. Advanced encryption concepts, such as attribute-based encryption (ABE) [BSW07], move the decryption capabilities from a specific (fixed) recipient to recipients with specific properties (attributes). To improve usability, Ma et al. [MHK+18] proposed an enhanced encryption scheme especially targeted for the industrial context that is able to make encrypted information searchable. Traditionally, systems processing solely encrypted data must rely on additional indexing schemes to support search queries based on this extra information.

*Data usage control* [PHB06] is another concept in the area and helps to enable information sovereignty. It allows distributing decisions regarding data access to multiple

| | authenticity | integrity | accountability | auditing and verifiability | immutability and referenceability | confidentiality | granularity | authentication | authorization | information control | identifiability | untraceability |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Data Security** | | | | | | | | | | | | |
| Encryption | | | | | | ++ | | | | | | |
| Data Usage Control | | | ++ | + | | ++ | + | + | + | + | | |
| Secret Sharing | + | | ++ | | | + | | | | ++ | + | + |
| **Data Processing** | | | | | | | | | | | | |
| Secure Offloading | + | | | | | ++ | | | + | ++ | + | + |
| Secure Computation | | | −− | − | − | ++ | | | ++ | | | |
| Verifiable Computing | + | | ++ | ++ | ++ | + | | | | | ++ | |
| Anonymization | − | − | − | | −− | ++ | ++ | | | | ++ | + |
| **Proving Support** | | | | | | | | | | | | |
| Digital Fingerprints | ++ | | ++ | | + | | | | | | | −− |
| Digital Signatures | ++ | ++ | ++ | + | | | | | | | | −− |
| Distributed Ledgers | ++ | ++ | ++ | ++ | + | | | + | | | | +/− |
| Version Control | | + | + | ++ | ++ | | | ++ | | | + | − |
| **Platform Capabilities** | | | | | | | | | | | | |
| Access Control | | | | + | | ++ | + | ++ | + | + | | −− |
| Policies | | | | + | | ++ | ++ | ++ | ++ | +/− | | |
| Smart Contracts | + | + | ++ | + | + | | | + | + | | | +/− |
| Confidential Computing | ++ | ++ | + | + | | ++ | + | + | + | ++ | + | + |
| Federated Learning | | | | | ++ | + | | | | − | | |
| **External Measures** | | | | | | | | | | | | |
| Data Markets | | ++ | | | | +/− | + | ++ | ++ | + | + | − |
| Legal Contracts | ++ | | ++ | + | + | − | ++ | ++ | + | ++ | ++ | −− |
| Smart Payments | | | | ++ | ++ | | | | | + | +/− | +/− |

**Table 2.1** A mapping between our surveyed building blocks (y-axis) and our categorization of relevant properties (x-axis) shows that no single one fits all solution exists. Depending on the security goal, the applicability of the different building blocks also varies significantly (from $++$ over $+$ and $+/-$ to $-$ and $--$). No entry denotes that no direct impact is notable.

parties. Hence, this approach fulfills all aspects of the challenge pertaining to the scope of information access. With the correct set of policies, logging functionality to achieve accountability can be integrated as well. Even though data usage control is more a (theoretical) concept than an established functional system, research in the area gained traction in light of information sharing in the IIoT [ZMJ+19].

Finally, *secret sharing* [Sha79, Ped92, Sta96] allows data sharing with multiple entities in a confidential way. To reveal the information, a subset of the entities must jointly reconstruct the original information, ensuring a certain degree of information control. Hence, apart from computational overhead, its applicability might be limited in dynamic industrial environments where entities often change. Regardless, Zhou et al.[ZC11] show an application in the IoT to establish a security architecture. In a more static context, Cyran [Cyr18] uses secret sharing in another domain (healthcare) with strict confidentiality requirements. Related approaches could help to overcome any trust issues companies in the industrial landscape might still have.

**Data Processing.** This category covers approaches that try to hide information during computations from unintended recipients, i.e., they extend the concept of simply limiting access to information to approaches that can also operate on or with it in a secure manner. In particular, we identify four groups of approaches: *secure offloading* [Che16] (operating directly on ciphertext), *secure computation* [MR91] (jointly computing a function without revealing individual inputs), *verifiable computing* [DSB17] (maintaining verifiable results even when offloading computations), and *anonymization* [SCDF16] (a collection of one-way functions to anonymize data).

Specific implementations of *secure offloading* support different computing complexities, e.g., homomorphic and order-preserving encryption (cf. Section 2.3.1). They have in common that encrypted data is sent to another party who performs calculations on the ciphertexts without inferring the content. Entities with the correct key can then decrypt the resulting ciphertext(s) to obtain the result(s). Recently, spooky encryption [DHRW16] (where encrypted inputs result in a plaintext result after computation) emerged to further improve the usefulness and applicability of secure offloading in practice. Such approaches enable stakeholders to rely on (untrusted) cloud services for computation without the fear of leaking information [ZPH+17], i.e., confidentiality and information control are preserved. Furthermore, it allows stakeholders to offload their computations anonymously because no conclusions about the data owner can be drawn. The IIoT is a prime applicant as companies operate with large amounts of process data and other business-related information.

Approaches in the area of *secure computation*, such as secure multi-party computation [Lin05], oblivious transfers and private set intersection (cf. Section 2.3.1), and zero-knowledge proofs [GMW91], provide protocols for multiple (distrusting) stakeholders to jointly compute a result or to obliviously exchange information (secrets). Hence, they are particularly suitable for information flows among previously-unaffiliated collaborators. Related work [DDM+19] even demonstrates privacy-preserving database lookups without the need for a trusted third party to reduce any leakage. Unfortunately, being oblivious reduces the accountability and referenceability of this approach significantly because the individually-provided inputs are only locally available. Hence, no (external) verification is possible without cooperation.

*Verifiable computing* [DSB17] addresses this shortcoming by ensuring verifiable results even when offloading computations to (untrusted) entities. To achieve public or administrative verifiability (i.e., properties that express who can verify a claim), it builds on secure computations, most importantly on zero-knowledge proofs. For widespread deployment of verifiable computing in the industrial landscape, we postulate a stronger focus on how long claims remain verifiable and whether they are forwardable [Len22]. Consequently, the corresponding participation and traceable verifiability properties require more prominent dissemination in research.

Fourth, various *anonymization* concepts, such as k-anonymity [Swe02], differential privacy [DR14], data aggregation [HLN+07, SCR+11], and noise [DKM+06, GN08], allow entities to protect sensitive information by aggregating it with other data points or by altering their precision. Then, they can collaborate with other stakeholders without leaking sensitive information. While these techniques also limit the accountability and authenticity of information, they potentially allow stakeholders to participate anonymously as no single data point can be traced back to a single entity. For example, in the IoT, related work showed that consumer usage data can be properly anonymized [HIFZ17, HPH+17]. This direction is likewise promising in the IIoT, e.g., to enable anonymous comparisons of the efficiency of production processes across manufacturers [PGH+19]. However, companies should take into account that information flows can already reveal relationships among different entities based on communication patterns only [HPD+19].

**Proving Support.** Moving from building blocks on access control, we now focus on the authenticity of information. Respective approaches range from proving physical aspects of a workpiece, i.e., digital fingerprints [VKW+16, PAAD18], to providing evidence for the origin and correctness of digital information (e.g., digital signatures [RSA78], distributed ledgers [MWM+16], and version control [LM12]). While different in scope, these approaches have in common that their ability to attest the authenticity and integrity of information contradicts the desire of stakeholders to remain untrackable. *Digital fingerprints* of physical products, i.e., having a unique digital identifier of a workpiece or product available, are difficult to realize in industry because attaching a barcode or a unique identifier to a manufactured product is not always possible. Consequently, new solutions are required to reliably link products to digital information to prove their authenticity and to remain accountable.

Nowadays, *digital signatures* are commonly used in the context of the Internet to provide authenticity, and this concept can be extended easily to the IIoT to provide similar verifiability there. *Distributed ledgers* have proven to be a suitable approach to improve auditing and immutability capabilities of this traditional solution. Blockchain technology (cf. Section 2.3.3) allows for establishing a persistent record of information and past information flows. Thus, they are a good fit for environments where multi-hop traceability (along the supply chain) is a strict requirement. Similarly, *version control systems*, such as Git, are also suitable for tracking data changes and enabling audits. These properties are required when dealing with a global knowledge system like the IIoT. However, in contrast to distributed ledgers, they are not tamperproof. Moreover, current version control systems might not support industry-specific data formats and volumes without adjustments or overhead.

**Platform Capabilities.** Apart from the technical building blocks encountered so far, we can also apply mechanisms that define and enforce rules for industrial collaborations. On the one hand, the traditional idea of *access control* [SS94] can help to restrict the scope of information access by setting rules for all individual entities. However, such restrictions are only possible if the participating entities can be tracked. Here, approaches from the IoT [LXC12] could potentially be transferred to the industrial sector. On the other hand, *policies* [HE02] directly attached to the data or exchanged information can offer similar flexibility [HHS+16] because usage or access constraints are efficiently retrievable (cf. data usage control). Instead of defining access rules for each entity, policies constrain the scenarios where and how a specific piece of information can be used, i.e., they are independent of the entity processing the data giving some control to the data owner, i.e., the company.

From a different perspective, the concept of *smart contracts* [Woo14,CD16,GLD+18] links the idea of blockchain technology with the benefits of automated contracts. Consequently, apart from proving the authenticity of information, smart contracts are also able to enforce the scope of information access to a certain extent. Previous work already applied them in the IoT [ZKS+18]. However, the implications of flexible and dynamic relationships on this design remain an open research question.

Additionally, *confidential computing* realizes an isolated enclave where guarantees about the running code and, thereby, about data accesses can be made (cf. Section 2.3.2). However, when using such enclaves, any interaction with the secure environment, i.e., incoming and outgoing information flows, still requires careful analyses. Besides, confidential computing could hinder interoperability as well as redeployments. Moreover, it might result in vendor lock-ins because specific confidential computing variants must be chosen. Furthermore, past security issues [CCX+19, FYDX21] showed that mitigating threats might require new hardware instead of simply deploying software patches. In the IIoT, replacing all computing hardware with confidential computing is highly unlikely as legacy devices frequently remain in use for decades [SPL+15]. Hence, corresponding (security) upgrades and deployments must be planned meticulously. Still, prior work [PGP+17] demonstrated the feasibility in an industrial context while addressing relevant security properties.

Finally, *federated learning* [LSTS20] pushes machine learning in decentralized settings and deployments. As such, it allows participating entities to keep parts of their sensitive information (including inputs) confidential while still benefiting from globally-available knowledge and information. Even though privacy and reliability implications are still open research questions, especially in the context of the IIoT, the overall concept promises significant (security) benefits for all involved stakeholders. Corresponding federated learning applications could even allow for process and scheduling adjustments based on external information [PHW21].

**External Measures.** The last category of building blocks contains supporting approaches that might help realize secure industrial collaborations. To monetize the value of sensitive information in the IIoT, (distributed) *data markets* enable all participating collaborators to sell and buy access to information. Besides mediating access to information, such a central data market can also ensure authentication and authorization, i.e., provide features that are relevant for the access to data.

Depending on the exact implementation, confidentiality can also be ensured if data is only shared in an encrypted format. Recent examples, such as the IDS [Int19] or other data markets [MMZ+17], have shown that centralized concepts to securely share information are feasible even in larger contexts. However, such a centralized approach shifts a lot of information to this market place which is, in turn, a valuable target for attackers of industry data. Thus, it might even turn into a measure that impedes industrial collaborations.

A less technical approach to restrict the scope of information access and to establish authenticity of information could be to rely on *legal contracts* [AW08]. They allow for defining all kinds of requirements before initiating the first flow of information. However, such negotiations are not yet automated in any way and, therefore, might prove infeasible in a highly-dynamic industrial landscape. Regardless, the concept can be used to set a frame in which entities are willing to collaborate and then negotiate the exact parameters in an automated way. Even though such an implementation would also allow entities to define sanctions in case of misbehavior, monitoring their actions and identifying data leaks from a remote vantage point is extremely challenging. Consequently, this building block might only be applicable to information flows in long-lasting business relationships.

To still facilitate automated information flows, stakeholders could also make use of *smart payments* [KL18a], which allow them to automatically initiate data transfers once the recipient has instructed a payment. As smart payments are likely based on distributed ledger technology, they also likely support auditing. However, they do not deal with securing the information and data access, i.e., instead of securing existing information flows, they make new dynamic information flows accessible.

### Takeaways

The scope and results of our survey underline that no one-fits-all solution that addresses all relevant properties is available. Instead, stakeholders must select the appropriate technology in accordance with their use case and security needs. Then, they can restrict their collaborations and information flows to settings and stakeholders that match their standards. So far, individual building blocks are only suitable for a small subset of the identified information flow challenges. Especially, well-founded research in the direction of industrial needs of confidentiality, verifiability, and anonymity is still in its infancy, resulting in insufficient coverage for highly-sophisticated use cases and real-world deployments. We consider the potential of transferring and applying established approaches from other domains to the IIoT as significant. While many interesting use cases require evolved building blocks to secure industrial collaborations, little progress has been made. Recently, related work [GSU+22] picked up our survey in the context of data markets. In this dissertation, we make a first step in this direction by addressing this shortcoming. In particular, we detail how to realize collaborations securely and scalably using technical building blocks that are well-established in (traditional) information security.

Before detailing our contributions (Chapters 4 and 5), in the next chapter, we first give an overview of selected use cases from the domain of production technology.

# 3

# Use Cases

In the previous chapter, we provided a structured overview of industrial collaborations and which building blocks might be appropriate to secure them. In this chapter, we move beyond this theoretical view and take a look at practical, real-world use cases that source corresponding information flows. Accordingly, in Section 3.1, we discuss the different types of collaborations in the industrial landscape in light of specific applications. Subsequently, in Section 3.2, we present several real-world use cases from the Cluster of Excellence "Internet of Production". We made a diverse representative selection to cover a wide range of settings with largely-differing (security) requirements. Ultimately, these sections conclude the combined and structured presentation of background information in this dissertation.

## 3.1 Industrial Collaborations in Practice

Following the introduction of the entities and different types of industrial collaboration (Section 2.1.2), we now look at the benefits of two entities collaborating more closely. Novel information flows promise improvements that are otherwise either not realizable or not as easily accomplishable. Our presentation follows the same order as Section 2.1.2.2, with the collaborations having been illustrated in Figure 2.1.

#### Type ①: Supplier and Manufacturer

In the industrial landscape, we identify multiple forms of collaboration among suppliers and manufacturers. We detail their specific applications in the following.

**Product Supplier.** The fundamental idea is that any supplier shares properties of the supplied items along with expected usage properties to enable reliable adjustments of the running process. These supplied items can be materials, parts, goods,

intermediate products, tools, or even production machines. Hence, the extent of digital information attached to such items might vary significantly. For example, a supplier of rolled metal can provide additional product details and properties of its production to the manufacturer of fine-blanked components [NKU+20, OBNB23]. In the opposite direction, the manufacturer can transfer details about the expected usage requirements for any ordered item. This information helps the supplier to only deliver items that fit the intended use. For example, a lower hardness of a non-structurally used metal piece might help to reduce the machine's wear without impacting the product. However, these details are sensitive as they can reveal process details to the supplier.

**Machine Supplier.** When considering machine suppliers, even more information might be exchanged. For example, the manufacturer could continuously share details about machine downtimes as well as about used components of the machine with the machine supplier. In return, the machine supplier could share estimates on the machine's downtimes and various condition changes with the manufacturer. This approach could even be turned into a Manufacturing-as-a-Service (MaaS) business model where the manufacturer only rents capacity on a machine from the supplier. Then, the machine supplier is responsible for all maintenance-related tasks and keeping the machine ready for operation. However, in this setting, the machine supplier has direct access to the production process, which reveals sensitive details of the production. Apart from theoretically being able to reverse-engineer aspects of classified manufacturing processes, the traceability of productive and non-productive periods of individual production machines can become an issue because they reveal a lot of sensitive information. By exploiting this internal knowledge of the manufacturer, machine suppliers could maintain a better position in future sales negotiations.

**Tool Supplier.** Similar observations on the scope of received information also hold for tool suppliers that provide, for example, cutters, cutting inserts, or grinding discs. The goal of sharing information among a manufacturer and a tool supplier is to reduce downtimes by accessing production-specific data to provide tools on time. For example, a tool supplier for milling machines should replace worn milling cutters on time with the most suitable model. However, in contrast to the machine supplier, the tool supplier might not have direct access to the production process. Hence, the respective access to production parameters is limited. Nonetheless, specific process parameters or wear characteristics can also reveal information about running processes and handled material (e.g., aerospace-grade aluminum for military vehicles), allowing the supplier to obtain knowledge about manufactured products, scheduled maintenance times, or the utilization of the production site.

**Security Perspective.** For supplier-manufacturer collaborations in general, we identify two primary challenges. First, sharing any information might allow the reverse-engineering of products or production processes and could even facilitate thefts of intellectual property or business secrets. This challenge affects both entities alike: Supplier knowledge can also inadvertently flow from the manufacturer to the supplier's competitors, for example, if (unmetered) data or machine access is in place. Today, legal contracts are commonly negotiated and signed to address such issues. Unfortunately, lengthy contract negotiations might not be applicable to supply chains

with short-term business relationships that lack any kind of established trust. In addition to the potentially infeasible overhead of setting up detailed contracts for a single business transaction, business relations without mutual trust might entail significant concerns regarding improper utilization of exchanged information or information leaks in general. Thus, the open question of how this practice translates to an evolved industrial landscape with its highly-dynamic and flexible business relationships and collaborations remains. The second challenge is that the involved entities could deliberately deliver incorrect values to yield (monetary) benefits, e.g., to artificially decrease the lifetime of a product or tool with the intention of boosting repeated sales. The situation intensifies when information further propagates along the supply chain, greatly impairing the reliability of exchanged information.

#### Type ②: Maintenance Provider and Manufacturer

Collaborations of this type are similar to the previous type, given that the maintenance provider also interacts with the manufacturer. Hence, corresponding information flows, such as usage values in one direction and usage estimates in the other direction, are very similar, which also translates to the security perspective. Typically, maintenance providers directly establish collaborations with manufacturers to offer their services, i.e., without the involvement of the original manufacturer of the machine. As a result, the maintenance provider not only receives valuable information about the collaborating manufacturer but, potentially, also about the manufacturer of the maintained machine. In the other direction, firmware updates or configuration recommendations might be passed along from the machine manufacturer or the machine supplier via the maintenance provider to the manufacturer to limit the number of entities that are directly involved with the machine.

**Security Perspective.** Given its direct access, the maintenance provider can gain valuable insights into processes and products of the manufacturer. However, business relationships of the maintenance provider with direct competitors of the manufacturer pose an even bigger risk concerning unintentional transfers of knowledge among clients of the same maintenance provider. In addition, manufacturers should verify the authenticity of updates or configuration settings that are passed along from the maintenance provider to prevent any misconduct. Likewise, maintenance providers want to avoid any liability claims following repairs or recommendations.

#### Type ③: Manufacturer and (End) Customer

We identify two conceptual information flows among customers and manufacturers in the industrial landscape. First, the customer can receive maintenance recommendations and firmware updates from the manufacturer (if provided) to improve the product's availability and productivity. Second, the customer shares her usage requirements along with usage values to send feedback to the manufacturer. This kind of information flow is identical to the relationship of a supplier and a manufacturer as the point of view defines the role within the supply chain (cf. Figure 2.1), i.e., the customer can also be another manufacturer that purchases manufactured products

from a supplier (the manufacturer in our point of view). Hence, in the following, we focus on customers in the role of a merchant or an end customer instead.

In this situation, customers can usually gain less sensitive (and interesting) information from the manufacturer. Thus, the interests for establishing such collaborations are imbalanced. Accordingly, manufacturers can offer their end customers discounts or other benefits in exchange for providing sensitive (usage) details. This information can, for example, support the manufacturer to link customer satisfaction as well as wear with particular subcomponents of a product, i.e., identifying responsible suppliers for contributing to exceptional or poor performance. Furthermore, detailed usage data can help the manufacturer to provide improved support to her customers, increasing both her knowledge about the product and the customer's satisfaction. For example, some machine tool manufacturers provide a process ramp-up service, where they support customers in finding stable process parameters for new machining processes in exchange for knowledge about the products being machined on their machines. Their motivation is to further increase their process parametrization expertise and secure their business interests and customer loyalty.

**Security Perspective.** The risks of tracking and surveillance based on usage information threaten end customers in such collaborations [YXSW18]. Without proper anonymization or aggregation of information, other supply chain actors might be able to identify customers based on backward-shared information. On the one hand, this threat especially emerges for products with only a few highly-specialized buyers as their privacy is particularly at risk. On the other hand, uncertainty about the reliability of received information threatens the involved manufacturers. With only a low number of customers, such usage data can have a significant impact on the decision-making of product development, procurement, and manufacturing. For example, they might falsely adjust their product based on incorrect usage data, which can ultimately have a negative performance on future production batches.

### Type ④: Manufacturer and Collaborator

The evolved industrial landscape will also feature collaborations across supply chains, e.g., with competitors or companies utilizing the same type of machines, tools, or components. Still, the corresponding information flows can vary significantly. For example, both entities might receive materials, parts, goods, or intermediate products from the same supplier. In this case, they would have an incentive to exchange knowledge about how to process the received items in the most beneficial way. In another scenario, both entities could operate machines or tools by the same manufacturer. To utilize them efficiently, they have an incentive to also rely on external experience, e.g., which settings reduce the machine wear or by sharing the key performance indicators about achievable output rates. Such "collaborations" are even conceivable among competitors if both stakeholders have identified benefits. If permitted by law, multiple companies can also establish a syndicated procurement. Collaborations across supply chains promise significant benefits simply by making information and knowledge globally available and utilizing it.

**Security Perspective.** Such collaborations are highly relevant for two reasons. First, they are crucial for the success of an evolved industrial landscape as they promise significant advances in the IIoT (cf. Section 1.2.1). However, stakeholders have significant security and privacy needs when dealing with collaborations across supply chains. Such that most potential following globally-shared information will remain untapped without secure and scalable realizations. Second, due to the flexibility of collaborating entities, questions of trust and accountability are especially challenging. Nowadays, long-lasting business relationships do not call for sophisticated and revisited concepts as signed contracts regulate most rights and duties anyway. However, when dealing with dynamic and short-lived business relationships, lengthy contract negotiations are infeasible. Instead, the industrial landscape requires technical means to address the concerns of involved stakeholders. In particular, aspects of accountability and reliability require careful consideration, especially when collaborating with anonymous entities. Here, security measures must ensure that collaborations do not negatively affect (participating) honest parties.

After this general overview of industrial collaborations in an (evolved) industrial landscape, we focus on representative real-world use cases in the following.

## 3.2 Representative Use Cases for Collaborations

To illustrate the aforementioned types of industrial collaborations, we now introduce selected use cases that originate from the Cluster of Excellence "Internet of Production". They are representative for our research in the IIoT as they cover all four types of collaborations. In particular, in Sections 3.2.1 and 3.2.2, we detail two use cases with collaborations along supply chains that cover Types ① and ③. As such, we also consider them representative for Type ② (as we have argued in Section 3.1). In Sections 3.2.3 and 3.2.4, we then focus on two use cases that benefit from collaborations across supply chains (Type ④). This way, we provide a comprehensive foundation for the presentation of our contributions in Chapter 4.

### 3.2.1 Product Composition and Production Properties

Our first use case represents manufacturing processes that benefit from information about previous production steps. For example, companies could alter their production processes based on information received from their suppliers. Moreover, they could consider substituting or combining certain parts or components to react to varying qualities of intermediate products. Such information could even be passed over multiple hops, e.g., to allow customers to better estimate the product's lifetime. Overall, the application of product and process information along supply chains offers various advantages. In case the received information is reliable and does not reveal sensitive information, companies could broadly apply and utilize it. In the following, we illustrate this use case at the example of two specific applications.

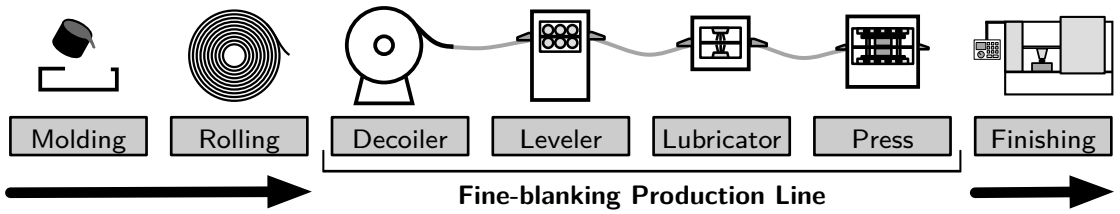



**Figure 3.1** Fine-blanked products are greatly influenced by the different production steps.

### Fine-Blanking Production Line

Our first application of this use case in industry is a fine-blanking production line. Fine blanking is a precision forming process that is applied to manufacture large batches of identical workpieces, e.g., for aerospace and automotive production [KK09, GHW+19], because it offers an excellent quality of the sheared surface and geometric accuracy [ZZZ19]. Maintaining an identical product quality is challenging even with a static production line setup [VTF+18]. In Figure 3.1, we illustrate the components of a fine-blanking line and its previous production steps, such as molding and rolling the input metal coil. The complex interplay of the decoiler, leveler, lubricator, and press reacts quickly to fluctuating material properties (e.g., as observed in coils), changing environmental conditions, or alternating behavior of components in a fine-blanking line [VTF+18]. Thus, detailed information on the supplied material and its production processes, even over multiple hops, is highly relevant, e.g., to adjust the parametrization of local production lines accordingly.

In an evolved industrial landscape, manufacturers might even benefit from collaborations across supply chains, i.e., Type ④. Thus, operators of fine-blanking lines could profit from directly exchanging process know-how to reduce scrap. At the same time, they could share the properties of fine-blanked components with their customers along the supply chain to inform them about minor quality deviations, which could potentially influence the customers' processes. Exchanging information along the supply chain could improve fault detection in case of failures because additional information to pinpoint the root cause would be available. Accordingly, all components have to be traceable and must allow for definite linking between analog products and digital information.

### Assembly of an Electric Vehicle

Looking at another application of this use case, we consider the assembly of an urban electric vehicle. In this competitive industry sector, many changes to the design and production processes of traditional vehicles were needed to ensure profitability. Given that the costs for the electric battery make up a large fraction of the overall production, the assembly of such low-cost electric vehicles requires a critical analysis of the supply chains of all remaining components. For a visualization of the entire supply chain in all its complexity with every involved component, we refer to our previous paper [BPM+21, Figures 6 and 7]. Since such an electric vehicle consists of 90 pre-assembled components with varying complexity, the number of

previous production steps and involved stakeholders (suppliers) differ significantly across components. Hence, an electric vehicle sources a large network of suppliers.

Due to large numbers of involved suppliers (also over multiple hops), the assembly of electric vehicles depends on extensive information flows to ensure that all safety and quality requirements are fulfilled. With so many involved stakeholders, all collaborations in this setting need to facilitate accountability. Otherwise, precise attributions of responsibilities, liabilities, and contact persons are impossible to implement, which, in turn, impairs the strictly-needed industrial collaborations.

With their large networks of suppliers and several previous production steps, the outlined applications of this use case concern multiple stakeholders and further demonstrate the complexity of products and production processes in the IIoT. These applications are representative for collaborations along supply chains and focus on Types ① and ③. To also consider more specialized production processes in the IIoT, we also select a use case on the operation and procurement of machine tools.

### 3.2.2 Operation and Procurement of Machine Tools

Machine tools are known for their challenging configuration, their regular maintenance schedules, and the complexity of running a stable production process. Due to these aspects, we consider machine tools as part of our second representative use case. In particular, milling machines are a prominent machine tool example. Given their complexity and variety, they are an important research area for engineers. However, we leave corresponding (technical) challenges for related and future work. Instead, in this dissertation, we look into their operation and procurement as industrial collaborations are likely to significantly impact both tasks. While the former task can benefit from sourcing information from suppliers and maintenance providers (e.g., to improve PPC, both on the shopfloor and within their network of suppliers), the latter is a prime application to further digitalize and automate business processes along supply chains, for example, by privacy-preservingly comparing whether the buyer's and seller's price expectations match.

#### Operating Connected Job Shops in Discrete Manufacturing

Discrete manufacturing outputs distinct products, i.e., products are manufactured separately. Consequently, ensuring consistent quality is important as defective intermediate products can have a cascading impact, both locally and globally. Apart from subsequent production or assembly steps, out-of-tolerance products greatly influence customers and suppliers because they might have to adjust their production processes and schedules accordingly. Unfortunately, without in-depth details, identifying and tracking down the root causes of inconsistent product qualities or impacted product lifetimes is tedious, time-consuming, and oftentimes also difficult. A larger information source (potentially even from collaborations across supply chains), including additional information, such as the machining tool condition and history, the processed workpiece material, and known deviations of the machine tool, promises to ease this task. Hence, this application benefits from information flows.

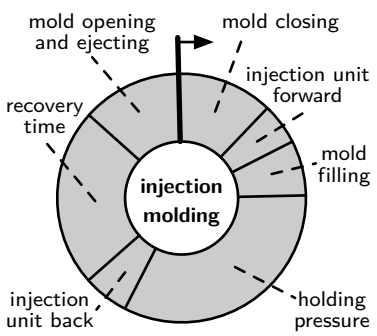
Connected job shops are embedded in global supply chain networks. For example, a manufacturer of milling machines integrates various received parts (drive and guide components, bearings, and milling spindle) of its supplier. Likewise, it contracts maintenance providers to replenish required tools, such as milling cutters, to both itself and its customers. Information flows among all involved stakeholders could positively impact the estimated product lifetimes (and guarantees), the load within production schedules, and the efficiency during product development. Even more, manufacturers could better react toward change requests as additional historical knowledge is available in a processable form. Such an evolution would most likely contribute to the continued success and prevalence of digital factories. First, MaaS business models could eliminate market barriers because their availability reduces the costs to build on machine tools for production. Second, traditional shopfloors and modern line-less mobile assembly systems (LMASs) [BMW+21, KWBK+23] (as popular in production sites of electric vehicles) can swiftly react to schedule changes.

### Procurement with Machine Tool Suppliers

In addition to the operation of machine tools, their procurement is another important application. Especially when looking for new products or variants, manufacturers have the incentive to look for new suppliers because (a) they could offer lower costs and/or better quality in comparison to their existing suppliers or (b) their existing network of suppliers might not be able to supply the requested item at all. However, without any contractual relationship, both parties have to exchange sensitive information that could reveal details of their products, production schedules, and new developments. This issue is particularly critical when ordering machine tools as they are highly individualized for the model (for example, in the automotive domain) and, thus, might reveal sensitive information on technological advances [Xu17].

While the manufacturer (buyer) does not want to disclose anything that might reveal aspects of her future model(s) or technological advances that are incorporated in the machine tool, the supplier (seller) does not want to reveal her entire product catalog and capabilities. Without secure collaborations, this situation leads to a dilemma: The more accurate the information, the better the offer; however, at the same time, the more accurate the information, the more sensitive details are being exchanged. The outcome of this dilemma is twofold. Either buyers do not contact potential suppliers at all, or they only contact a few suppliers with existing trust relationships or contracts. This pre-selection severely restricts the market and excludes a number of potentially more suitable suppliers, greatly impacting their businesses.

To elaborate, manufacturers tend to establishing close partnerships with only a few trusted suppliers (single sourcing) [YTR20], which leads to a strong increase in dependency [Ind08]. Even more, such a strategy severely impacts future adjustments in terms of quality, functionality, innovation, and costs [Ind08]. Contrarily, other buyers risk the “leaking” (disclosing) of information by diversifying broadly. To balance those two extrema, many manufacturers rely on indirect methods, such as signing non-disclosure agreements (NDAs) early on [WR17]. However, such non-technical approaches do not ensure confidentially with malicious-but-cautious stakeholders (cf.




**(a)** Injection molding process

| Dimension | Organizational | | Technological | |
|---|---|---|---|---|
| General Capabilities | Customer Satisfaction | | Manufacturing Processes | |
| | Labor | Finance | Resources | Prod. Range |
| Strategy-specific Capabilities | Flexibility of Due Dates | Mold Construction and Development | Technology Lead Process | |
| | Development Delays | | System Competence | |
| | Internationalization | | Technology Lead Part | |

**(b)** Benchmark perspectives in the injection molding industry

**Figure 3.2** Benchmarks capture complex processes and their operation in meaningful KPIs.

Section 2.1.2.1). Thus, technical means that evolve procurement processes would be highly beneficial as machine tools with their sensitive information exemplify.

In the context of machine tools, this use case shows the benefits collaborations along the supply chain and with maintenance providers can have (Types ①–③). However, they also stress the confidentiality concerns (and needs) stakeholders might have in an evolved industrial landscape. Moving toward collaborations across supply chains and their needs, we next look at the use case of company benchmarking.

### 3.2.3 Internal and External Company Benchmarking

As a first use case that covers collaborations across supply chains (Type ④), we consider company benchmarking. We distinguish two types of benchmarks with varying stakeholders: internal (measuring departments of a single company) and external benchmarking (comparing multiple companies) [Koz04]. While benchmarks could reveal sensitive information to "competitors" (regardless of being internal or external), they lack any direct influences on processes or production schedules. Thus, this use case is limited in its invasiveness and a good first step to pursue.

#### Benchmarking Operations in the Injection Molding Industry

Our first example covers an external benchmark from 2014 that focuses on the injection molding industry, i.e., it measures the efficiency of the injection molding department. Injection molding is widely applicable in different industries and domains and allows for the processing of complex part geometries without subsequent rework. As we detail in Figure 3.2a, it is a highly-complex discontinuous process and consists of various production steps. The raw plastic material is plasticized by heat and friction and then injected into the mold, which is the negative of the plastic part to be produced. After a pre-defined cooling time, the final part can be ejected from the mold [KIL09]. Accordingly, the real-world benchmark covers an organizational and technological perspective, as we illustrate in Figure 3.2b. While organizational key performance indicators (KPIs) capture the financial status of a company as well as the satisfaction of customers and employees, technological KPIs benchmark the

efficiency of manufacturing processes (such as the productivity of machines), means of production, and range of the manufactured products.

Given the lack of secure industrial collaborations, it relied on extensive manual labor by the analyst (operator of the benchmark). In addition to sharing questionnaires with queries to the participants, the analyst has to digitize the participants' answers. The analyst then inputs this digitized information into the self-developed benchmarking algorithm to output the KPIs, which are subsequently shared with the respective participant. Thereby, the analyst is a valuable target as all sensitive information is retrievable from a single point. For example, KPIs representing process information (e.g., quality of the manufacturing processes) can be sensitive as competitors could derive or estimate the participant's status and ongoing evolution.

In addition to the questionnaire, which contained 423 distinct questions and collected a total of 674 inputs per participant to eventually compute 48 KPIs, the complexity of the benchmarking algorithm (more than 2700 computations per participant) is also significant. Thus, we also have to note the analyst's effort. Given that the algorithm constitutes the analyst's intellectual property and competitive advantage, this application demands confidentiality of the algorithm in addition to ensuring the participants' privacy. Accordingly, operations with paper-based questionnaires and an analyst with access to all confidential information are not timely for the IIoT.

### Measuring the Efficiency of Global Production Networks

In addition to injection molding, we also consider a second benchmarking application that measures the performance of production sites in globalized production networks [Rit18]. This diversity allows us to study the use case in a different setting and with another kind of algorithm. In this context, benchmarking the performance of individual production sites is particularly interesting to compare the efficiency of companies or locations within a single company. For example, a KPI can express the unit costs of a product at a specific location for this purpose. By breaking down the unit product costs, companies can then identify the main drivers, such as the degree of automation, the wage level, or even the characteristics of the machine park. While production networks can also be measured as part of an internal benchmark, we focus on external benchmarks due to their stricter confidentiality needs.

Distributing production sites and supply chains can yield significant advantages as such an organization allows stakeholders to optimally exploit each location's geographic, regulatory, and technological conditions [VMPL21]. However, a competitive advantage is only present if the assumed-to-be-beneficial performance is verified regularly. To assess their performance, companies compare their inventory, efficiency, and equipment across different days, products, and orders in such benchmarks. Previous work [EGBH15,HHF+20] has identified the product portfolio's complexity as a major driver of costs, which can also be traced back to the need to interrupt production sequences with setup processes. This need results in reduced machine utilization and drives up unit costs. By nature, this information is highly sensitive, and thus, it must be kept confidential. Otherwise, competitors could draw conclusions about the company's corporate strategy and relationships [KOWFC10].

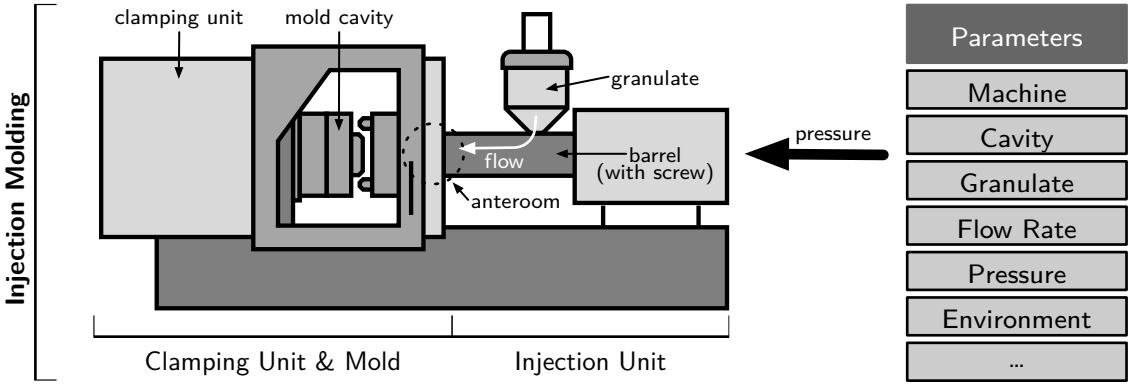



**Figure 3.3** Injection molding is a discontinuous process that is influenced by various parameters.

To summarize, this use case is a good first step to demonstrate and promote secure collaborations across supply chains in an evolved industrial landscape. Given its diverse security needs, it allows to effectively stress the benefits and the justified deployment of technical designs to various stakeholders. Still, company benchmarking is not too invasive to scare conservative stakeholders away. Moving on, we broaden the scope (and safety implications) of collaborations across supply chains by looking at the use case of sharing sensitive business parameters in the following.

### 3.2.4 Sharing and Exchanging Production Parameters

Finally, our fourth use case covers the sharing and exchanging of product and process details among manufacturers across supply chains. Once such collaborations are available privacy-preservingly, companies could adapt their operations according to global knowledge to reduce costs, improve product quality, reduce wear, and operate more sustainably. Especially for manufacturers with production lines that are costly and time-consuming to set up, such as injection molding, the expected benefits are significant. In addition to this first application, we further look at the potential for machine tools in connected job shops as part of our second application.

#### Commissioning and Configuring Injection Molding Production Lines

As outlined in Section 3.2.3, injection molding is a discontinuous production process. In Figure 3.3, we illustrate an injection molding machine with the injection unit and the clamping unit to toggle the pressure in the mold cavity. Polymer granulate is fed into a barrel that is heated using heater bands from the outside of the barrel. A rotating screw within and a translational drawback motion create friction through the interplay of granulate, melt, screw, and barrel surfaces. Friction and heat cause the material to plasticize along the axial transport to the screw's tip, which fills the anteroom with melted granulate. Eventually, the machine injects the accumulated melted polymer under high pressure into the cavity of an actively-cooled mold. Once the cavity is filled, the machine must press material into the cavity to compensate for volumetric shrinkage, which occurs during the cooling process. This step is also

known as the pressure-defined packing phase. Finally, the machine can eject the solidified workpiece [OTG07], which concludes the injection molding process.

This complex process features various production parameters that require adjustments according to the respective machine and cavity, the inserted granulate, and environmental conditions, among others. Hence, when commissioning new production lines (or configuring for different products), a lot of scrap is being produced simply by grid testing different configuration parameters. Here, the clever utilization of experience from other stakeholders could likely help to accelerate this costly and time-consuming task. However, given the sensitivity of the relevant production parameters, they may not be publicly available, i.e., suitable designs must account for such privacy needs when establishing corresponding information flows.

Apart from injection molding, including high-pressure die casting [PGH+19], with its trial-and-error configuration approach, other processes with the need for manual tuning could benefit as well. In the context of this dissertation, fine-blanking lines [PHS+19] and machine tools, which we briefly discuss next, come to mind.

### Operating Connected Job Shops in Discrete Manufacturing

Applying the use case of exchanging production parameters across supply chains is also reasonable for machine tools. For subtractive manufacturing (e.g., turning and milling), the workpiece quality and productivity of the machine greatly depend on the choice of configured cutting parameters, such as cutting speed, feed rate, and cutting depths. Without the utilization of global knowledge, these parameters are usually determined based on experience or manufacturer-specific recommendations. Both approaches are time-consuming tasks as the operator explores and fine-tunes the parameters' performance by repeatedly manufacturing workpieces.

Tapping into globally-available experience is a promising approach to obtaining optimal cutting parameters for certain requirements, such as roughness and tool lifetime [MST+14]. Enriching this utilization with a model-based approach and real-time process data even allows for estimating achievable optimizations of the configuration parameters [BWW19]. To conclude, connected job shops with machine tools are complex setups that depend on highly-accurate machining. However, trial-and-error approaches for optimizing cutting parameters are costly and time-consuming in this application as well. Thus, sourcing information from manufacturers across supply chains would constitute a valuable addition as accurately modeling machine tools is difficult and not always feasible [PBL+20].

Securely realizing collaborations across supply chains (Type ④) would allow stakeholders to operate their machines more efficiently while remaining in control of all processes. With increasingly-available knowledge, the usefulness of exchanged information will likely further increase over time. Even though the benefits are manifold, the acceptance of such collaborations still depends on their security guarantees.

This overview of our use cases concludes this dissertation's presentation of fundamental background information. With our use cases in mind, in the next chapter, we detail our designs and contributions for collaborations along supply chains.

# 4

# Collaborations Along Supply Chains

In this chapter, we look at (secure) industrial collaborations along supply chains. In particular, we consider two different settings. As part of our first contribution, in Section 4.1, we primarily focus on information flows that follow from established business relations, i.e., the collaborating entities know each other, at least locally. Accordingly, we also look at improvements for current best practices that follow from secured industrial collaborations. Second, in Section 4.2, we abandon this premise and specifically study the challenge of finding and bootstrapping new suppliers for business relationships along the supply chain as part of the procurement process. In this second contribution, we explicitly consider the confidentiality needs in settings with unknown, most likely untrusted entities. To the best of our knowledge, we are the first to improve the privacy guarantees of this essential task.

## 4.1 A Processing Pipeline for Reliable Information

In our first contribution, we focus on existing business relationships along the supply chain. As we have outlined in Section 2.1.2.2, corresponding information flows are beneficial for various goals (cf. Section 1.1). However, given the sensitivity of the processed and shared information, we have to carefully consider the trade-off between transparency and confidentiality when implementing collaborations and information flows along supply chains. Moreover, as several benefits are only achievable when adapting local production processes and schedules, the reliability of received information is of utmost importance for all involved stakeholders.

Moving on, in Section 4.1.1, we first introduce the flow of sensed and processed information along supply chains. Subsequently, we cover the challenges of reliable sensing (Section 4.1.2) and privacy-preserving information sharing (Section 4.1.3). We refer to the combination of both aspects as our processing pipeline. Finally, in Section 4.1.4, we conclude the presentation of our first contribution (said pipeline).

### 4.1.1 Concept of our Sensing and Information-Sharing Pipeline

As a foundation for our work, in Section 4.1.1.1, we first give a broad overview of digitalized supply chains, their logical actors, and common use cases that utilize supply chain information. These details are essential to understand the reliability and confidentiality needs of involved stakeholders. Subsequently, in Section 4.1.1.2, we present related work to capture previous efforts concerning information sharing in supply chains. We conclude this subsection with a high-level introduction of our processing pipeline in Section 4.1.1.3. Afterward, we discuss our designs to realize reliable sensing and privacy-preserving information sharing in supply chains.

#### 4.1.1.1 Information Processing in Supply Chains

In the following, we briefly recap our previous introduction to supply chains from Section 2.1. We further augment this presentation with details that are relevant when focusing on supply chains from an information-processing perspective.

##### Scenario Overview

When looking at supply chains from an information-processing angle, supply chains consist of a physical dimension (the flow of shipments, parcels, products, and paper-based documentation) and a digital one, which solely covers the exchange of information along the supply chain. Accordingly, we consider monetary flows as out of scope for our work. The digital dimension covers process, product, logistics, and scheduling (e.g., for PPC) information alike. All information is usually (i) acquired or sensed, (ii) transmitted or forwarded, (iii) processed, and (iv) eventually persisted somewhere. Depending on the usefulness and established information flows, (v) persisted information might be queried by various stakeholders. Hence, it is accessed to share it along, both upstream and downstream, the supply chain. In this scheme, updating past or outdated information constitutes a specific processing operation.

**Involved Stakeholders.** Most importantly, supply chains are usually composed of multiple stakeholders that are related to the product lifecycle (cf. Figure 1.1). They (can) range from excavation companies, over refining companies, suppliers, and manufacturers to customers and consumers. Moreover, distributors, retailers, commodity corporations, and many other entities can be part of a supply chain network. In light of developments toward a more sustainable IIoT, additional entities, such as recycling companies, are likely to become relevant. In addition to these *product-specific* stakeholders, the physical dimension also introduces *logistics-related* stakeholders, such as shipping companies, customs authorities, and warehousing services, as well as *sales-oriented* stakeholders, e.g., distributors, retailers, and customers. In digitalized supply chains, all of them are involved in the acquiring or sensing, forwarding, processing, persisting, and sharing of information. While the "simple" acquiring of information (i.e., accessing and sharing persisted information) is the most prevalent information flow along supply chains, our following descriptions specifically focus on information flows that start with a sensing step because this step raises several issues

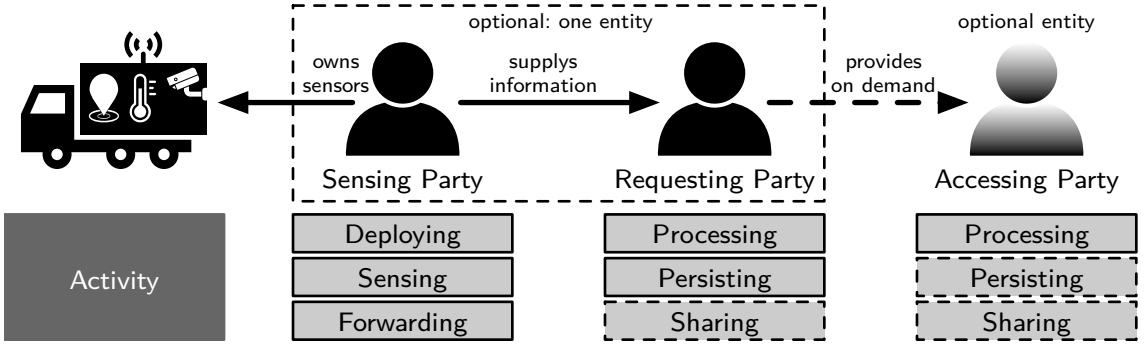



**Figure 4.1** We identify *up to* three logical actors for the sensing perspective in supply chains. Depending on the setting, some information-processing activities are optional (dashed boxes).

related to the authenticity of information. Depending on their information needs and use cases (see at the end of this subsubsection), a multitude of (sensed) information can be demanded and provided (shared) by individual stakeholders. Especially with indirect business relationships, i.e., over multiple hops, pre- and succeeding stakeholders might not be trusted or even known. Consequently, in such settings and when dealing with sensitive information, stakeholders require secure collaborations to ensure privacy-preserving and reliable information flows along supply chains.

#### Sensing in Supply Chains: Logical Actors in Information Flows

From an information-flow perspective, we deal with three conceptual actors, namely, sensing, requesting, and accessing parties, who are involved in the lifecycle of processed information along supply chains, as we illustrate along with their relations in Figure 4.1. Stakeholders along the supply chain can take the roles of multiple logical actors simultaneously, i.e., the logical mapping depends on the information flow.

First, a *sensing party* owns and deploys the sensors in use, either to capture information during the transit of products or to locally record information on products and production processes. It also makes sure to forward the sensed information to the requesting party. The sensing party is usually a shipment provider, but for shipments with sensitive, expensive, or fragile cargo, a customer might also request the inclusion of its own sensors. Likewise, warehousing departments can act as sensing parties when utilizing smart readers to process incoming or outgoing shipments.

Second, *requesting parties* are the intended, original recipients of sensed information that is forwarded by sensing parties. Requesting parties are, for example, direct customers who request (product or production) details on their purchase or the transit of their goods. The requesting party may also be the sensing party (for documentation or benchmarking purposes). Still, sensed information is usually initially meant for and requested by a single party only. However, if shipment providers group cargo by different customers in a single container, all customers might be interested, for example, in maintaining the cold chain. Looking at the availability of sensed information, after processing, the requesting party is responsible for long-term storage.

Third, additional *accessing parties* might be interested in (sensed) information (at a later point in time). Accessing parties can be virtually any stakeholders that

are concerned with the supply chain, from production companies, over suppliers, retailers, and governmental agencies (e.g., customs authorities), to end customers. Specifically, the requesting party initially shares the sensed information, and if no direct business relationship between the accessing and requesting party exists, information must pass multiple hops. In exceptional situations, e.g., following an accident, sensed information might only become relevant after years. Decoupled from the sensing, acquiring information is a broad step: Accessing parties can also directly retrieve information or processed derivations from other accessing parties, creating intransparent and long-lasting information flows. Such information flows entail confidentiality concerns because shared information can be highly sensitive.

### (Long-Term) Use Cases of Information in Supply Chains

As we detail in our survey on information flows in supply chains [PMK+24], various use cases source information (flows) to improve the supply chain performance and its management. In addition to the collaborative planning, the design of supply chains, the handling of shipments as part of critical infrastructures, and providing digital product information (together with shipped physical products), stakeholders are primarily concerned with the use cases of tracking and tracing in supply chains. Especially in light of dynamic and short-lived business relationships (cf. Section 1.1), the importance of these use cases increases further to ensure the accountability of involved stakeholders. While tracking primarily entails updates regarding the location of shipments and condition monitoring during transit, multi-hop tracing is essential when dealing with (origin) certification, production issues, or faulty products.

Given the multitude of tracking and tracing definitions, we now introduce these terms in more detail. In this dissertation, we follow the elaborate presentation of our survey [PMK+24], which slightly differs from the definitions in our previous work [PBM+20,BPM+21]. The differences are as follows: (i) tracking ($\rightarrow$) [PBM+20, BPM+21] referred to following product flows *downstream*, (ii) tracing ($\leftarrow$) [PBM+20, BPM+21] referred to following product flows *upstream*. Now, (iii) tracing ($\rightarrow$ and $\leftarrow$) [PMK+24] in this dissertation is bidirectional and encompasses both previous definitions, i.e., following product flows *downstream* and *upstream*. We distinguish these scenarios with the terms "handling faults" and "sourcing faults". In contrast to our previous definition, (iv) in this dissertation, tracking [PMK+24] does not correspond to following the product flow but is instead linked to a specific product.

**Tracking Use Case: Sensing Information.** Stakeholders in a supply chain are interested in information about upstream and downstream activities [GKHD20]. Accurate tracking information is of utmost importance for the management of supply chains, for example, to improve delivery date predictions and time slot management [PWB+19, BHFL20]. Given that various independent parties are involved in the production and shipment of ordered goods, uncertainties regarding compliance with (delivery) schedules and shipment environments arise. Accordingly, concerning the use case of tracking, companies are interested in two kinds of (sensed) information: (i) tracking data, i.e., where the shipment is, and (ii) monitoring data, i.e., what condition the shipment is in.

**Tracing Use Case: Sharing Information.** In contrast to tracking, the use case of tracing captures multiple dimensions. In particular, when handling faults, information flows downstream; when sourcing faults, information flows upstream. Moreover, as part of validating products and goods, information might flow in both directions of the supply chain. We now look into the different dimensions in more detail.

*Handling Faults (→).* Tracing individual components downstream is highly relevant when identifying issues with products after the fact. Recalls of faulty or dangerous products, including medical products [ODJ+22], as well as tampered or spoiled food products are prime examples [GKHD20]. Generally, the ability to identify subsequent products is beneficial for quality assurance as companies can recall those products that are potentially affected by a faulty production charge of a specific product. Secure multi-hop information sharing promises to reduce follow-up costs and latencies until taken (safety) measures show their effect.

*Sourcing Faults (←).* The ability to identify the root cause of a problem [AB20] as well as potentially affected subsequent products is relevant for several scenarios. To this end, the physical flow of a product or component is traced upstream through the supply chain [RHS23]. Here, quality assurance [BOS+21] and minimizing harm to customers and other businesses are key examples. Given the upstream branching of (complex) supply chain networks, the revelation of full supply chain structures and a dedicated inspection of specific production paths is crucial in this context. Sourcing faults upstream and then handling faults downstream promises benefits to various stakeholders and customers in the supply chain network.

*Validation (↔).* As requested by customers or mandated by law (e.g., the supply chain act [Bun21]), stakeholders need to abide by regulations and contracts regarding their processed or shipped goods, for example, to prove that their products are ethically and sustainably sourced [SAC19]. Particularly, sales-related stakeholders are interested in the product's origin. In the context of pharmaceutical products [Dei05, GABAS22] or art [ReZIB21], genuine products are crucial to avoid dealing with counterfeit products. Thus, complete, unmodified, and accurate historical information about the activities and production processes in supply chains is needed. Consequently, supply chains must support the reliable tracing of products.

Overall, tracing allows for advanced multi-hop product analyses as well as sophisticated quality control [MKJ18, AB20]. Identifying production issues and determining affected products positively affects the trust of customers and business partners, allows for improved lifetime estimates, constrains the need for maintenance downtimes, and further reduces costs resulting from undetected product issues. In the context of this dissertation, we thus specifically consider scenarios where tracking and tracing information is shared over multiple hops. Accordingly, we also cover information flows where receiving or sharing stakeholders are not known or trusted. Therefore, we have to address and secure collaborations in complex trust relationships. Moreover, defunct or merged companies complicate information flows over multiple hops. Finally, the sensitivity of shared information and the side effect of tracking and tracing to reveal supply chain structures mandate appropriate access control mechanisms. Otherwise, unintended information revelation or leaks might negatively impact the competitiveness of companies [SS02].

#### 4.1.1.2 Related Work

The ongoing transition to digitalized supply chains with deepened information flows has sparked various research activities. However, due to the opaque structure of business relationships, (legal) requirements, and the handling of sensitive information, impacting the evolution of supply chains and their well-established practices on a large scale is a challenging endeavor. Consequently, to shape the industrial landscape, new approaches need to securely realize reliable information flows. The expected benefits motivate us to look into current trends and developments concerning information processing in supply chains, which we summarize in the following.

**Research in Digitalized Supply Chains.** Traditionally, companies mainly considered their local perspective and restricted any information sharing along the supply chain [SS02]. As a result, issues such as the bullwhip effect [MCdD07] surfaced. Generally, digitalized supply chains can help to improve the transparency between two parties using automated information flows. Improving the transparency of activities, e.g., through secure collaborations along supply chains, promises various benefits [LEG99, SS02]. While Attaran et al. [AA07] present a general evaluation of different viable collaboration models, Flynn et al. [FHZ10] study the performance of supply chain collaboration in single-hop collaborations, i.e., they neglect the potential of multi-hop transparency. However, the digitalization of supply chains is not restricted to sharing information. Instead, sensing and monitoring shipments, schedules, products, and processes are seen as potent enhancements to enable fine-granular adjustments of the PPC and logistics operations at various stakeholders [PHS+19].

Still, business-oriented research in the context of supply chains and their management considers a multitude of research directions [CDPS+18, SN19, FK21]. The closest overlap with computer science concerns the processing of information [PMK+24]. Most prominently, related work puts great emphasis on tracing [GKHD20]. However, so far, only a few approaches [MSBAU22a] even consider the security of information flows over multiple hops. Even sophisticated solutions, such as DECOUPLES [MEE19], require the participation of stakeholders when tracing products. Recent work [KL18b, WVK18] proposed to realize multi-hop tracing in supply chains with smart contracts. More generally, we refer to our previous work for a large-scale survey of proposed approaches in the area [BPM+21].

We particularly notice the prevalence of blockchain-based approaches [BPM+21], which is considered to be a key technology for the evolution of information flows in supply chains [GKHD20]. Thus, we now look into blockchain-backed approaches.

**Supply Chains and Blockchain Technology.** Blockchain technology appears to be a natural fit for supply chains as it offers verifiable and tamperproof storage without requiring a trusted third party [HP17, KHD17]. In this regard, Wüst et al. [WG18] provide guidelines on whether blockchain technology is appropriate for a specific scenario, and they consider supply chains as one of their scenarios. Achieving that no single entity controls all information is strongly desirable in a setting with distrustful parties as prevalent in complex supply chain networks. Blockchains can ease the recording of trade events as well as help to improve (global) verifiability and account-

ability [DDJ+20]. Other directions look into asset recording, e.g., to identify counterfeit products [For23], to enable (origin) certification [MDKJ19], or to promote fair trade [AM16]. These improvements are especially desirable for food supply chains, and they even provide interfaces for governmental oversight [MKJ18, MDKJ19].

Overall, these approaches have in common that they are usually challenged by two aspects. First, all actors must be known in advance to allow for appropriate access control. Second, even though blockchain-based approaches promise tamperproofness and accountability, they are also susceptible to pointlessly storing irrelevant data (the issue of "garbage-in, garbage-out" [PFCN22]). Still sourcing blockchain technology, other domains look into the improved intersection of the digital and physical world [DJP+19, PV20]. In this context, hardware-based security concepts have not yet been explored, but research at the intersection of blockchain technology and confidential computing is ongoing to extend the applications of blockchains.

**Blockchain Technology and Confidential Computing.** Microsoft proposed the TEE-backed Confidential Consortium Framework [RAA+19] that realizes a replicated ledger inside a TEE. Other approaches explore TEEs for various blockchain applications [TLK+18, CZK+19, MWS+19]. Today, the first real-world blockchain deployments secure their operation using TEEs [Mob17]. Apart from this intersection, confidential computing is an evolving research area that also focuses on the IoT. Thus, it is also of interest when securing collaborations in the IIoT (cf. Section 2.4).

**Confidential Computing.** As we have indicated in Section 2.3.2, various research projects and commercial products are available [MGDC+17, SMS+22]. While research attempts to secure the sensing of information without confidential computing in medical environments [LSLN+16], two TEE variants might be particularly appropriate for this task in the IIoT: (i) the lightweight TEE Sancus [NBM+17] and (ii) the commercial TrustZone [PS19], Other embedded security architectures [KSSV14, NER+19] might be equally suitable if they feature isolation and attestation.

**Takeaway.** Various researchers tackle challenges related to the processing and sharing of information along supply chains. Especially blockchain technology is frequently applied to establish and improve respective information flows, especially due to its decentralized nature. However, the digitalization of all aspects of supply chains is an ongoing research area, and uncertainties regarding privacy-preserving information sharing remain. Even though the benefits of transparency are well understood, research still insufficiently studies the impact of information sharing over multiple hops. Likewise, concerns regarding the confidentiality of sensitive information or the use of novel technologies seem apparent, as proposed approaches rarely end up in large-scale real-world deployments. Most notably, we believe that fine-granular access control and technical guarantees are crucial to overcoming these concerns.

Concerning the sensing of information, prior work [WG18] raised concerns about the insecure and unreliable sensing (manipulation of sensors) and forwarding of information in connection with supply chains and blockchain technology. This issue persists for all sorts of information along supply chains that is acquired after being processed and shared (multiple times) because the original source, i.e., sensed information, might have already been inauthentic and thus unreliable. Simply deploying

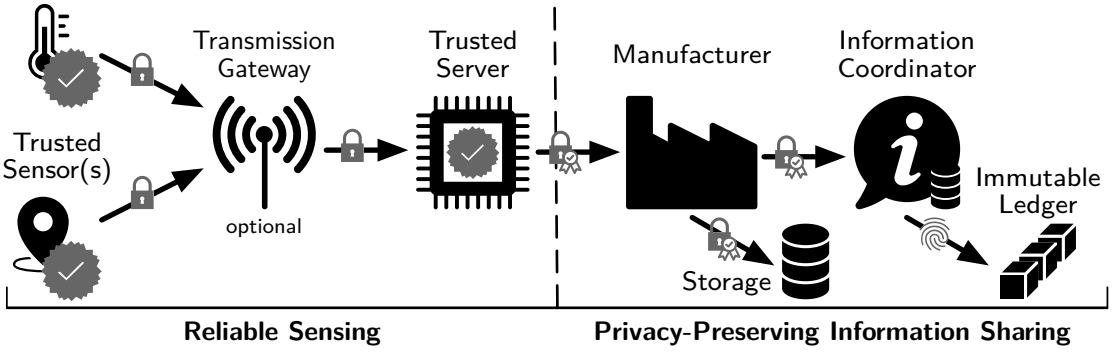



**Figure 4.2** Our processing pipeline for reliable information consists of two concepts. While the former ensures reliable sensing, the latter enables privacy-preserving information sharing.

IoT sensors is insufficient, especially in settings with business-oriented, mutually-distrustful stakeholders who might even have a (monetary) incentive to misbehave.

To the best of our knowledge, no prior work studied and resolved this matter using specially-secured sensors, i.e., confidential computing-backed hardware. Paired with the reliability issues following the widely-suggested use of blockchains (garbage-in, garbage-out), we identify the need for approaches that secure all steps of the information processing, from sensing to persisting (and sharing), in the IIoT.

#### 4.1.1.3 Overview of our Processing Pipeline

We propose a processing pipeline for reliable information in supply chains. As we illustrate in Figure 4.2, we separate it into two crucial challenges. First, with our "*Secure and Reliable (End-to-End) Sensing*" design, we focus on the sensing part. In particular, we utilize confidential computing (cf. Section 2.3.2) and trusted sensors (sensors that are secured using confidential computing). Second, our "*Long-Term Private (Multi-Hop) Information Sharing*" design ensures privacy-preserving information sharing along the supply chain. Inspired by related work, we ensure long-term verifiability of information by relying on blockchain technology (cf. Section 2.3.3). In the following, we present both designs in detail. The order of our presentation follows the flow of information, i.e., from sensing to persisting (and sharing).

### 4.1.2 Secure and Reliable (End-to-End) Sensing

In the IIoT, crucial information about shipments, processes, and products is sensed, forwarded, and processed by potentially-untrusted stakeholders. However, without reliable sensing (even in such remote environments), companies cannot utilize the information to the fullest extent. Accordingly, in this section, we introduce our design to enable reliable sensing in the IIoT. We make sure to propose a (i) cost-efficient, (ii) easily deployable, and (iii) maintainable design that guarantees authentic, untampered, and verifiable information. In particular, technical means ensure that companies can be held accountable (for their actions and forwarded information).

We refer to this part of our contribution as *reliable end-to-end (E2E) sensing*.

In the following, in Section 4.1.2.1, we first discuss the scenario and the research gap in more detail. After detailing the application of sensing in Section 4.1.2.2, we derive our design goals in Section 4.1.2.3. Subsequently, in Section 4.1.2.4, we briefly introduce the technical foundation of our design. Then, we present our design in Section 4.1.2.5 and evaluate it in Section 4.1.2.6. Finally, in Section 4.1.2.7, we conclude the presentation of the sensing part of our information-processing pipeline.

#### 4.1.2.1 Scenario Overview: Supply Chain Tracking and Monitoring

In light of our focus on sensing, we now extend our high-level scenario from Section 4.1.1.1. This level of detail allows us to accurately outline the research gap.

**Scenario.** As we have introduced (cf. Section 4.1.1.1), various stakeholders sense, forward, and process information along the supply chain. Additionally, manufacturers may locally record information on products and production processes for subsequent sharing with other stakeholders along the supply chain (both upstream and downstream). For a visualization of different shipment steps, i.e., tracking the shipment and monitoring its condition, we refer to our previous paper [PAM+20, Figure 1]. Accordingly, information frequently passes several (untrusted) entities before it is eventually processed, persisted, or utilized. Unfortunately, any entity with access to the taken communication path can compromise sensed (and shared) information. Consequently, given the lack of trust, companies require a technical approach that ensures the availability of authentic and untampered information.

Depending on the product or good, stakeholders are interested in different information. For example, in food supply chains, monitoring the upholding of a cold chain depends on temperature sensors. Additionally, humidity sensors can sense another angle of relevance. Other sensors can record any opening or closing of shipment containers or crates to verify that shipments have not been tampered with during transit. In general, the availability of various low-cost sensors allows companies to tune the monitoring granularity according to use case-specific needs.

**Research Gap.** Today's supply chains lack technical solutions that provide recipients of shipments and products with the means to rely on sensed information. However, numerous threats during the sensing, forwarding, and processing of information arise due to the large number of actors with potentially non-existing or low-trust relationships. As a result, stakeholders might apply inefficient processes instead (e.g., paper-based reports), decide to not collect desired information, or refrain from trusting received information. Thus, we identify the need to provide stakeholders with a secure E2E sensing design to allow them to reliably (trustworthy and timely) detect undesirable delivery statuses or environmental conditions of their shipments and other issues, even when untrusted stakeholders reported otherwise. Such a design then allows companies to attribute issues such as temporarily-interrupted cold chains beyond doubt. When pairing such a design with our privacy-preserving information sharing (cf. Section 4.1.3), we can also enable sophisticated long-term verifiability and accountability, e.g., to unequivocally assign blame in case of disputes.

#### 4.1.2.2 Prevalent Sensing Applications in Supply Chains

Based on the EPCIS standard [GS116,BWK17], we can identify a number of relevant sensing applications in supply chains. These applications all belong to the tracking use case that we introduced before (cf. Section 4.1.1.1). The corresponding EPCIS data model captures the dimensions of what, when, where, and why [GS116]; with the latter being irrelevant in our scenario. The upcoming successor, which has been designed for state-of-the-art supply chain data interoperability, further includes the dimension "how" [GS121]. This dimension matches the previously-outlined research gap of providing companies with trustworthy and verifiable information.

##### Sensing Applications

When dissecting the different types of sensing in supply chains, we identify five general applications with increasing (technical) complexity. These applications are not limited to a single granularity (e.g., shipments only). Instead, they can be applied on a shipment, parcel, or product level. By default, measurements include the dimensions *what* (i.e., the focus of the measurement) and *when* (i.e., timing information). Thus, in the following, we focus on the dimensions *where* and *how*.

*Status Tracking.* To track the status of a physical flow, requesting parties are interested in updates. For example, when handling a parcel, such as moving it from a container to a warehouse or changing the mode of transportation (e.g., from truck to aircraft), this information, as well as its location (*where*), must be recorded. This tracking can be achieved with stationary RFID readers or BLE beacons (*how*). Generally, this application is well-known from tracking consumer parcels.

*Location Tracking.* If requesting parties also want to know about the approximate locations of physical flows, they demand a periodical or real-time update of the respective locations. Thus, location sensors (*how*) must reliably sense this information (*where*). In the context of consumer parcels, this sensed location data is frequently available nowadays as part of tracking updates during last-mile deliveries.

*Integrity Monitoring.* Apart from these tracking applications, companies might also be interested in the (physical) integrity of their shipments (*where*). For example, when dealing with pharmaceuticals, detailed documentation might even be required by law [Foo18]. Potential violations can result not only in monetary damages but also in harm to humans (e.g., food poisoning). Thus, precisely capturing such data and maintaining access logs are important aspects. Correspondingly, sensing parties can deploy digitally-secured (smart) locks and other surveillance sensors (*how*).

*Condition Monitoring.* Extending the previous application to continual or real-time monitoring (*where*) can also be relevant. For example, compliance with temperature ranges is crucial in cold chains, e.g., to identify spoiled goods. Likewise, manufacturers might define constraints for shipment environment conditions (e.g., humidity or impacts). Thus, deployed sensors must reliably provide this information (*how*).

| Sensing Application | Sensor | Payload | Measurement Frequency | Time Criticality |
|---|---|---|---|---|
| Status Tracking | Smart Reader | <1–<100 KiB | Triggered | Minutes |
| Location Tracking | e.g., GPS | <1 KiB | <1 records/min | Minutes |
| Integrity Monitoring | Smart Lock | <1 KiB | Triggered | Hours |
| Condition Monitoring | Various | <1–<10 KiB | <6 records/min | Hours |
| Visual Monitoring | Camera | >10 KiB | Variable | Variable |

**Table 4.1** The sensing applications feature different sensing equipment and technical needs.

*Visual Monitoring.* When considering very valuable products or livestock, video-based monitoring (*where* and *how*) is an application as well. Depending on the setting, the corresponding image or video feed might only be transmitted after specific triggers, e.g., opening after unlocking a smart lock or when exceeding a specific noise threshold. Thus, the exact needs vary significantly in light of the specific goal.

Stakeholders commonly rely on a combination of these applications. Depending on the exact setting, several types of sensors with different densities must be deployed: Location, temperature, humidity, air pressure (altitude), light, shock (impact), acceleration, tilt, or weight sensors, as well as smart locks and scanners.

### Technical View on Sensing Applications in the Industrial Landscape

Due to the associated computational burden of these sensing applications, we now consider payload sizes and sensing frequencies (i.e., the processing bandwidth). Latency is of interest to guarantee a timely handling of status information or condition changes. We provide a summary of relevant parameters in Table 4.1.

Overall, most applications have moderate needs when sensing relevant information. The exact payload size depends on the sensor in use, as well as the size of added context information. For status tracking, payload sizes can range from as little as 96 bit for the most common type of RFID tags to as high as 100 KiB or more for extremely-specialized RFID tags with extended user data [WNYD09, CAP+13]. Similarly, location tracking and integrity monitoring can be realized with less than 1 KiB, most of the time. Different sensor types introduce varying payload sizes for condition monitoring, with simple measurements taking up only several bytes. In contrast, visual imagery monitoring or video feeds may lead to more excessive needs (far greater than tens of kilobytes), depending on the desired image quality and resolution, as well as the availability of potent compression methods.

The individual sensing frequency for condition monitoring can be comparably high (i.e., several measurements per minute). However, corresponding latency requirements are usually not demanding because condition monitoring is intended to serve as an authoritative reference, i.e., companies usually do not act upon updates during the transit of a shipment. Essentially, while the information gathered by status and location tracking may be needed within minutes of gathering, integrity and condition monitoring is likely to produce sensed data that tolerates (processing) delays of a few hours. Data aggregation at the source could help to lower the amount of data

to transmit, especially if requesting parties are only interested in outliers. Thus, in practice, the continual transmission of large payloads is unlikely in most settings.

#### Misbehavior: Threats to the Credibility and Reliability of Sensed Information

With multiple supply chain actors involved, we need to consider several types of misbehavior, as misbehaving actors could try to deceive requesting or accessing parties. Essentially, all misbehavior builds on one or multiple of the following actions.

*Data Tampering.* Dishonest parties might have an incentive to directly manipulate data during transmission for various reasons. Shipment providers could, for example, try to cover up deficiencies concerning the shipment's treatment. This desire may emerge in cases of accidents that should not reflect negatively on the company's reputation or in cases where requesting parties try to deceive accessing parties.

*Data Hiding.* If direct data tampering is not possible, simply hiding the existence of data or of a specific range of data may be equally desirable. The lack of data may not be surprising to the victim (honest actor) and could easily be blamed on unreliable technology or environmental events such as power outages.

*Data Injection.* A malicious party could also attempt to insert forged information, i.e., data that originates from unauthentic and unrelated sources (sensors) or is made up entirely. Similar to the previous misbehavior, such actions could be useful to deceive actors and convince them to accept the present shipment conditions.

With these sensing applications and threats of misbehavior in mind, we derive a set of corresponding general design goals in the following.

### 4.1.2.3 Design Goals for Secure and Reliable E2E Supply Chain Sensing

Establishing secure and reliable E2E sensing for real-world deployments depends on various properties. In the following, we introduce *five* indispensable aspects.

**G-S1: Tamperproofness.** Sensor data must be verifiably untampered when being assessed by the requesting or an accessing party at any point in time. This property is vital to ensure that sensor data can be relied upon by all parties.

**G-S2: Authenticity.** The design must ensure that sensed data verifiably originates from (the claimed) authentic sensors. Thereby, malicious actors are prevented from (retroactively) forging data. This goal covers both sensing and requesting parties.

**G-S3: Completeness.** To truly enable E2E-secured sensing, recipients (i.e., the requesting and all accessing parties) must be able to verify that the data they receive is complete while being able to unequivocally blame parties who are responsible for forwarding or sharing incomplete data. Apart from checking for the data completeness from a single sensor, recipients need to confirm that the measurements of all relevant (i.e., deployed and expected) sensors are available.

**G-S4: Latency Agnosticism.** Any solution must be agnostic to any network latencies or network disruptions (offline periods) experienced by the sensing nodes while generally supporting a wide range of applications, from frequent live updates to infrequent batch uploads.

**G-S5: Affordability.** Finally, given the overhead of any technical solution, both the (one-time) costs for additional hardware or hardware upgrades and the associated operating costs should be kept to a minimum. Consequently, new designs should be careful to (i) avoid performing computation-intensive tasks and (ii) not introduce excessive duty cycles during (regular) operation. Otherwise, real-world deployments are unrealistic on low-cost IoT devices, preventing widespread adoption in industry.

When a design fulfills the goals **G-S1** to **G-S3**, it also adds accountability to scenarios with untrusted actors as all information is complete and verifiably attributable to the sensing party. In connection with a design that offers (secure) long-term storage of (sensed) information, e.g., a blockchain recording all sensed information or PrivAccIChain (cf. Section 4.1.3), the outlined guarantees even hold long-term.

#### 4.1.2.4 Preliminaries: Lightweight Confidential Computing for the IIoT

We utilize a specific confidential computing architecture in our reliable sensing design. The relevant details exceed our previous introduction of confidential computing in Section 2.3.2. Thus, we now briefly highlight its most important features.

**Sancus.** Sancus is especially suited for IoT and IIoT deployments as it is built on the MSP-430 [NBM+17], a 16-bit processor family. Its computational resources are sufficient for basic IoT applications (Sancus runs at a speed of 20 MHz and provides 65 kB for code and data). The availability of *memory-mapped input/output (MMIO) modules* in the MSP-430 permits attaching sensors to the device. Sancus can then assign these MMIOs to a specific protected module that guarantees isolated sensor access [NMP17], i.e., its design ensures that the TEE has a secure, exclusive channel to the sensor readings. Like other variants, Sancus also supports secure communication channels (within the TEE). Consequently, it can guarantee a secure forwarding of sensed information to remote parties for persisting and further processing.

In the following, we refer to sensors that are backed by confidential computing (either using the Sancus security architecture or ARM TrustZone) as *trusted sensors* as their hardware is able to guarantee that their sensed information is (i) authentic, (ii) untampered, and (iii) verifiable.

#### 4.1.2.5 Reliable End-To-End Sensing in Supply Chains

Now, we present our design to improve the reliability of sensed information along supply chains using technical means. In particular, we rely on confidential computing to establish authentic and verifiable information flows of sensed data, essentially establishing a notion of end-to-end-secured sensing. After giving a brief design overview, we discuss the different components that are involved in our design in more detail.

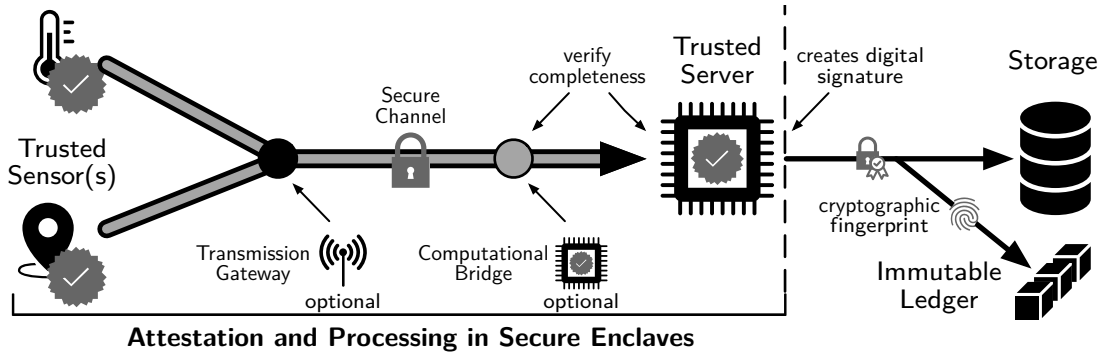



**Figure 4.3** Between sensor and server, information flows through a secure channel. TEE-enabled hardware enables attestation processes and thus also allows checking for completeness of data. Depending on the setting, on-path nodes with (computational bridge) or without (transmission gateway) TEE support may be deployed. The trusted server prepares the sensed information for reliable long-term storage by creating a digital signature. A cryptographic fingerprint of the sensed information ensures integrity and is persisted on an immutable ledger.

### Design Overview: Securing Information from Sensor to Storage

In Figure 4.3, we illustrate the information flow from the sensor (left) over the processing (center) to the storage (right). We follow this order when discussing our design, which builds on confidential computing to provide technical guarantees about sensed information. Our work relies on *trusted sensors* that reliably sense all interactions with or conditions of shipments along the supply chain, even in untrusted environments. These sensors forward sensed information either directly to a *trusted server* that also operates within a TEE or to a transmission gateway that eventually relays the sensed information to a trusted server as well. The gateway may even be untrusted as it does not perform any cryptographic operations. If information should be altered or filtered in any way before it reaches the trusted server, TEE-backed *computational bridges* may be deployed on path. The secure channel enables the trusted server to attest the authenticity (**G-S2**) and completeness (**G-S3**) of sensed information. The trusted server further creates a digital signature to establish verifiability and persists all information at a *storage provider*. Persisting a cryptographic fingerprint of the sensed information on an *immutable ledger*, which can be based on blockchain technology, provides (long-term) tamperproofness (**G-S1**). Moreover, these fingerprints also act as proof of existence (for sensed information).

With this design, we (i) ensure that sensed information originates from trusted sensors, (ii) all information is initially protected through TEEs and secure channels between them, and (iii) cryptographic fingerprints persisted on an immutable ledger enable verifiability and tamperproofness. Information only leaves the protected computing environments after the trusted server has created corresponding digital signatures. Thereby, we can provide verifiable E2E guarantees from the source (sensor) to the tamperproof storage (immutable ledger) where cryptographic fingerprints of the information are persisted. Thus, respectively-sensed information is reliable.

### Processing Steps within our End-To-End Sensing

As outlined, information flows across different components in our reliable E2E sensing design. We thus discuss their responsibilities and duties in the following.

**Trusted Sensors.** All sensor data is protected from tampering (**G-S1**) through TEE-enabled microcontrollers with protected MMIO modules, e.g., as offered by Sancus. RFID readers, state-logging locks, IoT sensors such as temperature or humidity sensors, or cameras are sensor types that are suitable for deployment and use in supply chains. To ensure authenticity (**G-S2**), the TEE of the trusted sensor also establishes an attested connection (secure channel) to the TEE of the trusted server.

**Transmission Gateway.** As an optional component, our design supports the integration of an (untrusted) transmission gateway that gathers sensed information from multiple sensors, batches it, and forwards it to the trusted server. As such, it can relieve the sensors from energy-intensive tasks and reduce the hardware costs of the deployed trusted sensors. For condition or visual monitoring, the gateway can also buffer messages to cope with offline phases during transit, thereby addressing **G-S4**.

The requirements of a TEE-enabled computational bridge are identical to the requirements of a trusted server. Thus, we omit its detailed description.

**Trusted Server.** Each company operates its own TEE-enabled trusted server, e.g., based on Intel SGX, that directly receives the sensed information via an attested and encrypted connection. Hence, we introduce technical guarantees that sensed information is authentic (**G-S2**) and complete (**G-S3**), even if it originated from an untrusted environment, as all information is processed inside TEEs. The attestation also facilitates the detection of inconsistencies, such as incomplete series of measurements. The trusted server's main task is to transform and persist attested sensor readings in a way that they remain verifiable in the future. To this end, it creates a *digital signature* of sensed information, persists all data at a storage provider, and records a corresponding cryptographic *fingerprint* on an immutable ledger.

**Storage and Immutable Ledger.** The sensed (and persisted) information remains verifiable after sharing as accessing parties can simply compare its fingerprint to the cryptographic fingerprint persisted on the immutable ledger. Related work [ZNP15] showed that cryptographic fingerprints offer reliable accountability. We hold companies responsible for ensuring reliable data retention. Given that we only publish fingerprints on the ledger, higher sensing frequencies and larger information volumes do not overload the immutable ledger. Moreover, by relying on a dedicated storage provider, we do not require any (public) availability of sensed information by default.

The persisting concludes the initial reliable and E2E-secured information flow, and any recorded information can be shared on demand. We further refer to our previous paper [PAB+24] for detailed explanations of why the different components of our design are needed and how to implement the relevant operations, messages, and computational operations, i.e., it also covers the computational complexity and security guarantees of simpler designs. As we later discuss in Section 4.1.4.2, once the information leaves the trusted server, we can also integrate more sophisticated (storage) designs to better cater to the needs of sophisticated industrial collaborations.

#### 4.1.2.6 Feasibility Study of our Reliable E2E Sensing

We have to study the feasibility of our design from three dimensions. First, our processing pipeline for reliable sensing in supply chains must handle all kinds of sensing applications, irrespective of the exact domain, shipment, or product. Second, in addition to the technical foundation (**G-S1**–**G-S4**), we also need to consider the costs and overheads of widely deploying such a design (**G-S5**). Third, our work must be robust against various attacks. Accordingly, we discuss its security (guarantees).

##### Satisfying the Performance and Scalability Requirements

Apart from the trusted sensors, all components of our design are logical entities only, i.e., we can easily scale their computational resources using common approaches. The overall performance of our reliable information-processing pipeline thus depends on the processing requirements and the data throughput on low-end sensing hardware.

**Sensing Equipment.** Depending on the sensing applications, we can scale the computational resources of the trusted sensor as needed: We distinguish between 16-bit trusted sensors, such as Sancus processors running at 8 MHz, and more powerful 32-bit sensors based, for example, on ARM Cortex processors with TrustZone. We can further support lightweight sensors with a computational or transmission gateway to offload computationally-intensive operations or resource-consuming communication.

*Gateway Equipment.* Transmission gateways may be particularly useful in scenarios with longer offline phases, addressing the goal of latency agnosticism (**G-S4**). In such deployments, the gateway's computing resources are secondary as it primarily requires sufficient storage to buffer sensed information until a network connection is restored, which may take up to weeks when considering long-distance shipments. In other cases, significantly fewer (storage) resources are needed. The exact choice depends on the number of sensors, the data size, and the sampling rate of forwarding sensors. Considering the previously-discussed applications, each application would only require a data storage ranging from $1\,\mathrm{KiB} \cdot 60 \cdot 24 = 1440\,\mathrm{KiB} \approx 2\,\mathrm{MiB}$ per day for small sensing applications like location tracking and up to $10\,\mathrm{KiB} \cdot 6 \cdot 60 \cdot 24 = 86\,400\,\mathrm{KiB} \approx 100\,\mathrm{MiB}$ per day for the larger applications. For visual monitoring of shipments, the requirements can become arbitrarily large, which is why they must be discussed on a per-case basis. Overall, we conclude that companies can easily scale their gateways upfront while serving multiple (trusted) sensors at a time.

**Cloud Resources: Trusted Server and Storage Provider.** Before we look into the specific computational needs, we ascertain that components in the cloud (including our trusted server and storage) allow for horizontal scalability: Despite conceptually being a single entity, arbitrarily many servers can cover the duties of a trusted server to match the required performance capabilities. Likewise, database systems supporting vertical and horizontal scalability are readily available [RB16]. Thus, these components allow for horizontal scalability and do not constitute a bottleneck.

When looking at the computational needs following the information flow from sensor (here: Sancus) to server, we observe that setting up our reliable sensing introduces a

computational overhead of 130 ms. Moreover, we measure the time for sensing and processing, i.e., the sensor sends the measurement to the server, which re-encrypts it, computes a fingerprint, and signs the processed information. Our evaluation over 100 runs shows that even large payloads of up to 10 KiB can be processed roughly 6 times a minute, where the sensor consumes the majority of computation time, and our Intel SGX-based trusted server only takes between 55 and 150 ms. Only sensing applications with larger payloads, such as high-resolution visual monitoring or specialized RFID tags with large unique identifiers of up to 100 KiB, demand more powerful microcontrollers like ARM Cortex-based devices.

For more details on these measurements, we refer to our previous paper [PAB+24]. The corresponding evaluation artifacts are publicly available [SrcC21, SrcC24a].

**Immutable Ledger.** Our design utilizes an immutable ledger to allow for long-term information integrity and consistency. Accordingly, storing cryptographic fingerprints of the sensed information and its digital signatures (generated by the trusted server) is sufficient for regular operation. Such a fingerprint can have a size of only 124 B [BPM+21]. However, larger fingerprints (longer hashes) are possible for even stronger security guarantees. As we detail in our previous work [BPM+21], the performance of typical blockchains, such as Quorum [CON20], is sufficient for our purpose. On a generic server with two XeonSilver 4116 CPUs and 196 GB of RAM, preparing a corresponding transaction that persists the fingerprint takes 5.50 ms. We measure a throughput of 741 TX/s at the ledger, which conforms with previous performance evaluations [BSKC18]. We can further improve the ledger's throughput by recording multiple fingerprints in a single transaction, relying on sidechains or sharding, or utilizing meta-fingerprints that cover multiple measurements [BCD+14, ZMR18, BPM+21]. Therefore, we conclude that the ledger's performance is negligible when being utilized for our information-processing pipeline.

This performance analysis expresses the feasibility and scalability of our E2E sensing: computationally, corresponding solutions are suitable for real-world deployments.

### Cost Analysis of Equipping (Existing) Supply Chains

We intended that our design primarily sources cheap TEE-backed hardware to remain affordable (**G-S5**) and, thus, widely deployable in today's industrial landscape.

A single standard shipping container usually contains several crates that ship multiple goods. Depending on the setting, trusted sensors can be installed in crates or containers. When considering our sensing applications (Table 4.1), most likely, less than five sensors are needed per crate, which can be attached to a single Sancus processor in most cases. Realistically, corresponding off-the-shelf IoT sensors are purchasable for approximately 10 € per crate. For more demanding applications, slightly more expensive sensors might be required. While Sancus is not commercially available (expected to be around tens of euros [PAB+24]), TrustZone-enabled ARM microcontrollers are available for under 100 €. Since TEEs do not impose specific hardware requirements on the sensors, existing sensors can likely remain in use when upgrading to our reliable sensing. The same holds for recent IoT devices, as many recently-purchased microcontrollers readily provide TrustZone [PS19].

A TEE-enabled computation gateway (which is essentially a mobile trusted server) could cover additional tasks, such as location tracking or scanning RFID tags when crates enter or leave the device's proximity. The price of such a gateway greatly depends on the communication medium in use (e.g., satellite communication vs. GSM or Wi-Fi). With ARM TrustZone, the total price of a container's hardware could accumulate to around 300 €. In extremely resource-consuming settings, computation gateways with Intel SGX might further increase the overall hardware costs.

Looking at the remaining components, on the one hand, we note that TEE-enabled cloud servers are commercially available at marginally-higher prices than common cloud infrastructures [Int23]. Moreover, a single server can simultaneously handle several shipments, trusted sensors, and gateways. Thus, they entail little cost overhead. On the other hand, the costs associated with the immutable ledger are difficult to express as several stakeholders are involved in its operation. Potentially, the trusted server could even handle this aspect using its unprotected (non-TEE) computing capabilities. Given the moderate throughput requirements of our reliable sensing (see above), we expect marginal cost overhead from the immutable ledger.

Following these observations, we conclude that both deployment costs and operational expenses (including expenses for required communication channels) are manageable, especially given the expected reliability benefits, and thus, our design demonstrates an affordable concept (**G-S5**).

#### Security Discussion

We reasonably assume the security of the concept of confidential computing (**G-S1**) and, thus, consider corresponding software vulnerabilities (e.g., [VBOM+19]) as orthogonal aspects of our security analysis. Likewise, we refer to prior work on how to (i) seamlessly reuse sensors [YMB+15, ADY+19], (ii) securely integrate lightweight processors without TEEs [MMH+15], or (iii) manage the deployment, storage, management, and revocation of key material in our setting [VBMP17].

Our design's security largely follows previous work [NMP17] that aims for strong security guarantees in heterogeneous TEE deployments. We introduce a new notion of E2E-secured sensing by persisting pointers (cryptographic fingerprints) to all relevant activity on an immutable ledger. Due to the combination of trusted sensors and TEEs and attestation, we know that fingerprints recorded on the immutable ledger originate from physical events that have been sensed. Since successful remote attestation uniquely binds the execution of an enclave to a trusted sensor, the authenticity of respective measurements is also guaranteed (**G-S2**). When information flows through our processing pipeline, it is only processed using TEEs, which also allows the components to determine the completeness of sensed information (**G-S3**). Moreover, before any information eventually leaves attested TEEs, the trusted server creates a digital signature to ensure the verifiability of the sensed and processed information. By immutably persisting fingerprints of the sensed information (contributing to **G-S1**), we then establish long-term verifiability.

**Misbehaving Actors.** In the following, we discuss security threats that might potentially follow from misbehaving actors (cf. Section 4.1.2.2) in real-world settings.

*Measurement Manipulation.* In our setting, data tampering corresponds to the manipulation of sensor measurements, e.g., to cover up issues. The utilization of trusted sensors with support for remote attestation enables unequivocal detection of such software manipulations. Additionally, the persisted fingerprints on the immutable ledger ensure that any manipulation of sensed information is detectable as well.

*Measurement Withholding.* Companies could attempt to hide data, given that direct data tampering (manipulation) is not possible. To this end, they could either withhold or delete measurements. Since each trusted sensor numbers its measurements, such withholding attacks are generally apparent from a gap in numberings.

*Retroactive Data Removal and Manipulation.* In addition to measurements being hidden, we also need to consider the threat of hiding the existence of sensors and their measurements. Again, the fingerprints stored on the immutable ledger serve as proofs of existence and integrity protection that reveal the deletion or manipulation of information, thus providing tamperproofness (**G-S1**) and completeness (**G-S3**).

*Data Forging and Replaying.* Finally, misbehaving parties could attempt to insert forged information, i.e., submit measurements that originate from unauthentic and unrelated sensors or are made up entirely (data injection). First, as our design builds on digital signatures, the origin and authenticity (**G-S2**) of sensed information are ensured. At the same time, a unique numbering of attested measurements per sensor prevents replay attacks, as duplicates would be apparent. Second, TEEs prevent data forging attacks as they ensure that signatures originating from a sensor can be solely created for measurements that also originate from the respective sensor.

This security discussion underlines the appropriateness of our design and stresses the suitability of confidential computing and blockchain technology for reliable sensing in supply chains. In addition to these threats to our conceptual design, we further need to discuss attacks on the boundaries of our E2E-secured sensing.

**Open Physical Attack Vectors of our Reliable Sensing.** To fully assess our design's feasibility for real-world deployments, we also have to consider attacks at the foundation of our (technical) reliability and E2E guarantees.

*Direct Physical Attacks.* Unfortunately, no technical solution can prevent direct physical tampering with the utilized sensors (regardless of being trusted or not). However, by deploying sensors with tamperproof physical isolation of the digital components or by physically hardening their equipment [AK97, KJJ99], actors can react to this kind of threat. We see these mitigations as possible but non-trivial future work.

*Sensor Registration Manipulation.* Apart from physical tampering, misbehaving shipment providers could also register multiple sensor sets per shipment and ensure that even if some sensors record issues with the shipment, carefully-placed backup sensors report the intended measurements (e.g., by shielding them from environmental conditions). Before handing over the shipment, the shipment provider could then remove the faulty sensors to only report the shielded sensors' measurements. Similarly, shipment providers could place sensors in a manipulated environment that differs from the intended one. To mitigate both issues, aligned with ongoing efforts in the IoT [BEM22], we suggest that future work investigates how to verify shipment

sensors by aggregating environmental conditions of nearby sensors or different sensor types and how to perform semantic checks on such data. A related approach could be examining reported information using measurements from geographically-nearby shipments or past shipments on the same route.

*Linking of Product and Measurement.* Our design is further challenged by the lack of solutions that can reliably interconnect and link the physical and the digital world [WG18]. Here, recent advances promise to strengthen the linking through various means [VKW+16,PAAD18,IBM23]. Modern marking approaches like molecular fingerprinting [SCC+17, CPR20] are one example that can help to uniquely identify a shipment. Actors along the supply chain can then rely on this additional metadata to verify the binding between shipped products and reported measurements. Given the magnitude of this mitigation strategy, we leave corresponding feasibility and affordability evaluations to determine its large-scale suitability for future work.

Overall, we argue that any real-world deployment would inherently require human decision-makers in the loop to make judgments and decisions based on the output of the technical domain. This addition is not only necessary to handle any misbehavior but also to accommodate potential technical failures of the E2E-secured sensing.

#### 4.1.2.7 Conclusion

This section concludes the presentation of the first part of our processing pipeline for reliable information along supply chains. In particular, we have presented the notion and concept of *reliable end-to-end (E2E) sensing*, which we derived from five general design goals. Corresponding designs are well-suited for environments, such as the industrial landscape and supply chains in particular, where stakeholders do not trust each other and any sensed information. By providing technical guarantees through the use of confidential computing (especially trusted sensors) and blockchain technology, we are able to address these concerns, even in flexible and highly-dynamic environments. Overall, we have discussed the feasibility of this approach in terms of performance, costs, and security. We look forward to evolutions of our work and first real-world deployments, which highlight the benefits for participating stakeholders.

### 4.1.3 Long-Term Private (Multi-Hop) Information Sharing

After our focus on the authenticity and reliability of sensed information along supply chains, we now broaden our view on information sharing in supply chains. Depending on the setting, shared information covers various aspects, such as product or process details, tracking or tracing data, and origin information, among others. Combining an approach that ensures authenticity and correctness of shared information, even in scenarios with short-lived and indirect (over multiple hops) business relationships, with reliable sensing promises to improve the usefulness of shared information. It can further ease the application of (federated) process mining that covers multiple supply chain actors [vdA21]. To date, information is only rarely shared over multiple hops of a supply chain due to its sensitive nature, i.e., (valuable) information is only

available (and retained) locally [DDJ+20]. Consequently, exhaustive information that captures the entire supply chain of a product and its origin is often missing.

To establish accountable-yet-confidential information flows along supply chains, we thus present a design that supports information sharing in a privacy-preserving and accountable manner. We refer to our design as *PrivAccIChain* (reads: privacy chain), which corresponds to **Priv**acy-preserving and **Acc**ountable multi-hop **I**nformation-sharing platform for supply **Chain**s. In light of an evolving industrial landscape, we particularly consider dynamic and short-lived business relations. That is, (a) we realize multi-hop sharing that is independent of the participation of specific companies, and (b) we utilize the technical building block of attribute-based encryption (ABE) to implement proper access control even in opaque supply chains. Thereby, we ensure the confidentiality of sensitive information while enforcing its availability.

We refer to this part of our contribution as *privacy-preserving information sharing*.

In the following, in Section 4.1.3.1, we first discuss the scenario and the research gap in more detail. Afterward, we derive design goals in Section 4.1.3.2. Subsequently, in Section 4.1.3.3, we briefly introduce the concept of ABE before presenting our design *PrivAccIChain*, in Section 4.1.3.4. Our evaluation covers four subsubsections: In Section 4.1.3.5, we first discuss the performance of the utilized building blocks. We continue with an evaluation of our design that covers two real-world scenarios, namely, the production of fine-blanked components and the assembly of an electric vehicle, in Sections 4.1.3.6 and 4.1.3.7, respectively. Then, in Section 4.1.3.8, we discuss the security of PrivAccIChain, which also concludes our evaluation. Finally, in Section 4.1.3.9, we conclude the sharing part of our information-processing pipeline.

#### 4.1.3.1 Information Sharing in Supply Chains

As a foundation for our design that facilitates information sharing in supply chains, we now extend our high-level scenario from Section 4.1.1.1. In particular, we discuss how to model supply chains and outline the implications of low-trust relationships.

##### Representing Supply Chains as Directed Acyclic Graphs (DAGs)

As we have discussed (cf. Figure 2.1 and Section 4.1.1.1), supply chains consist of multiple actors that fulfill various tasks in a product's lifecycle. While the business relationships of these actors are bidirectional (i.e., goods flow in one direction and payments in the other direction), product flows are always unidirectional. Hence, we can always model them as acyclic processes. Accordingly, we can transform a supply chain (network) into a DAG that represents the flow of products (including their composition) linearly over time. Nodes without incoming edges usually represent actors that excavate natural resources. Intermediate nodes process incoming goods (they can also merge several goods into a single product) and supply manufactured products to their customers. The last node without an outgoing edge captures the last user/consumer of a product. Edges denote a change in ownership. As such, they usually also capture some sort of shipment from one stakeholder to another.

In intermediate nodes, different product flows involving various stakeholders are combined by production processes, indicating a composition of intermediate products or parts. In theory, this linkage allows involved companies to determine products that utilize certain subcomponents as well as to identify those components that are used by a composed product by traversing the DAG for information retrieval. Accordingly, the DAG illustrates ownership transfers as edges. Hereby, we represent logistics-related services as separate production processes, i.e., we model each product flow with only two different actions: namely, a *produce* step (a node in the DAG) and a *trade* step (an edge in the DAG). In our design (cf. Section 4.1.3.4), these two actions also create the foundation for our proposed data record structure.

#### The Implications of Low-Trust Relationships for Information Sharing

Modern, digitalized supply chains frequently establish information flows among direct business partners. However, we rarely observe any information sharing over multiple hops, i.e., indirect business partners, due to fears of leaking data and losing control over sensitive information, which would affect their businesses. However, common supply chain use cases (cf. Section 4.1.1.1) mandate the sharing of information over multiple hops, for instance, to globally distribute tracking and tracing information [GKHD20]. This dilemma contributes to companies limiting their information sharing to long-term business relationships with well-established trust [SS02], which is incompatible with the evolving industrial landscape (cf. Section 1.1).

In light of the evolving industrial landscape and the IIoT, in our research, we thus explicitly consider information sharing in settings with short-term business relationships. In particular, the lack of long-term trust must be addressed. In such low-trust environments, traditional approaches are unsuitable as they require mutual trust, which is unrealistic in such flexible supply chain networks. Furthermore, corresponding designs must especially prevent unintended (horizontal) information spills across supply chains. To offer significant benefits through multi-hop information sharing in combination with low-trust business relationships that do not negatively impact the companies' competitiveness are needed. Apart from the best practice of restricting the revelation of information to a minimum (as required for the respective use case), sharing stakeholders should remain in control of granting access to their information.

### 4.1.3.2 Design Goals for Privacy-Preserving Multi-Hop Information Sharing

Any approach that intends to improve information sharing along supply chains must satisfy various aspects that directly follow from (i) the outlined use cases (cf. Section 4.1.1.1), (ii) the stakeholders' concerns regarding information sharing, and (iii) dynamic supply chain structures. As a prerequisite, we mandate that suitable approaches are able to model entire supply chains, products, and business relationships. Moreover, in light of an increase in (short-lived) low-trust relationships in the IIoT, information sharing should not build on established trust relationships. Otherwise, corresponding approaches cannot provide extensive benefits in dynamic supply chain networks. Finally, novel approaches must break today's multi-hop barriers to (better) support global use cases and industrial collaborations, as we have

outlined before (cf. Section 4.1.1.1). Tracing use cases to handle (→) and source (←) faults are of particular interest. Consequently, new approaches should reliably, i.e., privacy-preservingly and securely, enable multi-hop information flows along supply chains. Based on these expectations, we derive *six* indispensable design goals.

**G-P1: Accountability.** Designs for information sharing in supply chains are only trustworthy and reliable, especially for long-term use, if they record information in a *persistent* and *tamperproof* manner while *attributing the ownership*. Given the global nature of supply chains, suitable designs need to particularly consider the prevalence of indirect (multi-hop) business relations. Then, designs can potentially establish global accountability for shared information as long as they provide guarantees on the *existence and availability* of historical information, protection against manipulation (*integrity*), and mechanisms to prevent the spreading of misinformation. Further, we explicitly demand that involved parties can be held responsible in cases of identified misconduct, even in the absence of deliberately-trusted parties.

**G-P2: Verifiability.** As a precondition to achieve accountability (**G-P1**), shared information must be verifiable retrospectively by all involved parties. Consequently, it must also remain *available* and *untampered* for later use or incident investigations. In particular, supply chain actors do not only want to detect manipulations but also want to identify cheating or misbehaving parties to hold them accountable. In the long run, sophisticated designs could even realize a reliable reputation system based on shared information and discovered disputes.

**G-P3: Privacy Preservation.** While this goal opposes the transparency implied by **G-P1** and **G-P2**, the acceptance of new designs depends on their ability to ensure that business secrets remain private whenever possible as companies are generally cautious when providing (sensitive) information. Thus, privacy preservation is a central design goal. Accordingly, suitable privacy-preserving mechanisms, such as encryption and (fine-granular) access control, are needed. Otherwise, stakeholders are likely to isolate their information, i.e., preventing multi-hop information flows.

Every proposed approach must be able to balance the trade-off between transparency (accountability and verifiability over multiple hops) and confidentiality (privacy preservation) as needed to ensure its usefulness while maintaining its acceptance.

**G-P4: Security.** We assess security as equally important because it ensures the technical guarantees for privacy preservation (**G-P3**). In particular, companies should require as little trust as necessary in the design and other stakeholders when participating. Businesses need to be able to rely on the enforcement of the demanded confidentiality regarding both malicious insiders and external attackers. That is, suitable designs must provide appropriate security features that *prevent unintended extraction or manipulation*. We expect that the complexity of supply chains and their corresponding use cases also negatively influence the complexity of proposed approaches. Hence, security must be carefully addressed in every setting.

**G-P5: Scalability.** Even today, supply chains, their actors, and associated product flows are complex environments with the need for a multitude of information flows and constantly-increasing amounts of sensed, processed, and shared information.

The situation only amplifies with an evolved industrial landscape that increasingly features short-lived business relations. Accordingly, designs need to account for a tremendous amount of shared information in terms of volume, velocity, and variety. This goal is not limited to processing overhead but also takes storage requirements and availability needs (whether a party must interact/participate within a specific time frame) into account. Thus, they also affect the affordability of newly-proposed approaches. Ideally, the designs' performance is independent of the number of participants to allow for virtually-unlimited and arbitrarily-complex supply chains. Consequently, this goal greatly contributes to any design's overall feasibility.

**G-P6: Autonomy.** Lastly, approaches for the holistic sharing of information in supply chains must remain autonomous, i.e., manual interaction should only occur following exceptional events, for example, to resolve claims concerning misbehavior. Today, sophisticated use cases, such as tracing or multi-hop information sharing, incur significant manual effort, preventing their widespread adoption. We can only ensure reasonable scalability (**G-P5**) in evolved supply chains by reducing the necessity for such manual interactions (even in the presence of accountability and verifiability). Furthermore, designs should not depend on the participation of actors or their availability, as their presence is not guaranteed in volatile supply chains. Solely with sufficient autonomy, designs might be suitable for real-world deployments.

These goals express the functional requirements for privacy-preserving multi-hop information sharing in supply chains. When pairing such a design with an approach that ensures reliable sensing (cf. Section 4.1.2), and thus authenticity and correctness of information, the usefulness of shared information improves even further.

#### 4.1.3.3 Preliminaries: Access Policies for Settings with Multiple Recipients

The complexity and opaqueness of supply chains challenge the traditional approach of encrypting information for specific recipients, especially if recipients are not known upfront (i.e., when encrypting the information). Therefore, we utilize a more general approach called attribute-based encryption that detaches decryption capabilities from a specific individual while offering fine-grained access control.

**Attribute-Based Encryptions (ABEs).** This novel form of public-key encryption shifts access control from *who* may decrypt data items to *which properties* or *attributes* are required for legitimate data access [BSW07]. To this end, formulas define cryptographically-enforced access policies, and all entities that satisfy a formula can gain access to the encrypted information. Consequently, in contrast to traditional public-key cryptography, encrypted ciphertexts are not bound to the recipients of information.

Several variants and implementations with slightly-different features are available to choose from [ZDX+21, MSBAU22b]. For our setting, we use ciphertext-policy attribute-based encryption (CP-ABE), which links ciphertexts to a logical formula of attributes [BSW07]. It supports efficient one-to-many settings, even if recipients are unknown at the time of encryption. A party can only decrypt the ciphertext if it satisfies the linked formula with attributes in its possession. For example, policy

$A \wedge B$ is only satisfied if the party has access to both attributes $A$ and $B$. Even two colluding parties with only attribute $A$ respectively $B$ in their possession cannot decrypt the corresponding ciphertext, i.e., entities cannot join their attributes to satisfy additional policies. While a central authority traditionally assigns attributes, the control over attributes can also be distributed to multiple authorities, each responsible for a set of attributes without a central coordinator [LW11].

ABE has been applied to and proposed for various applications, such as broadcast encryption [BSW07], cloud storage [KL10], or mobile cloud computing [KKKM13]. Other applications cover, for example, sophisticated privacy protection in social networks [JMB11]. Similarly, in the context of the IoT, Zhang et al. [ZZD18] apply ABE to address potential data security concerns in smart health applications.

#### 4.1.3.4 PrivAccIChain: Multi-Hop Information Sharing in Supply Chains

We now present PrivAccIChain, our design to realize transparent multi-hop yet privacy-preserving information sharing. We particularly account for settings without established trust. Accordingly, PrivAccIChain also supports fine-granular access control and provides long-term accountability and verifiability. In addition to the basic ability to share information with (indirect) business partners, our design explicitly considers the use case of tracing ($\rightarrow$ and $\leftarrow$). After giving a brief design overview, we discuss the different components that are involved in our design as well as the process of providing, retrieving, and updating information in more detail.

We refer to our previous paper [BPM+21] for all implementational details of PrivAccIChain. In this dissertation, we focus on the concept and our design decisions.

##### Design Overview: Sharing Information in Low-Trust Environments

The privacy and reliability guarantees of PrivAccIChain build on three technologies: (i) *AES* to securely encrypt sensitive information, (ii) *attribute-based encryption* to fine-granularly protect AES keys with a many-to-many encryption scheme, and (iii) *blockchain technology* to reliably record all actions persistently in a distributed way. In the context of secure collaborations along supply chains, we can always associate information sharing with the flow of physical products or goods. Accordingly, pointers between different information records correspond to nodes and edges in a DAG that expresses the product flow, i.e., products are either consumed (combined) or exchanged among supply chain actors. Following this design decision, we only have to support two actions: (a) a *produce* operation, which combines (intermediate) goods into a newly-manufactured or assembled part, and (b) a *trade* operation, which tracks whenever an item is physically handed over to another supplier, the contractor, or a customer. To improve the scalability of our design while still supporting fine-granular product flows and produce operations, we also allow for a batching of goods along (selected) edges. PrivAccIChain further provides an update mechanism for recorded information while maintaining a (verifiable) versioning system.

Shareable information originates from actors in the supply chain (here: *manufacturer*). PrivAccIChain acknowledges that excessive point-to-point communication

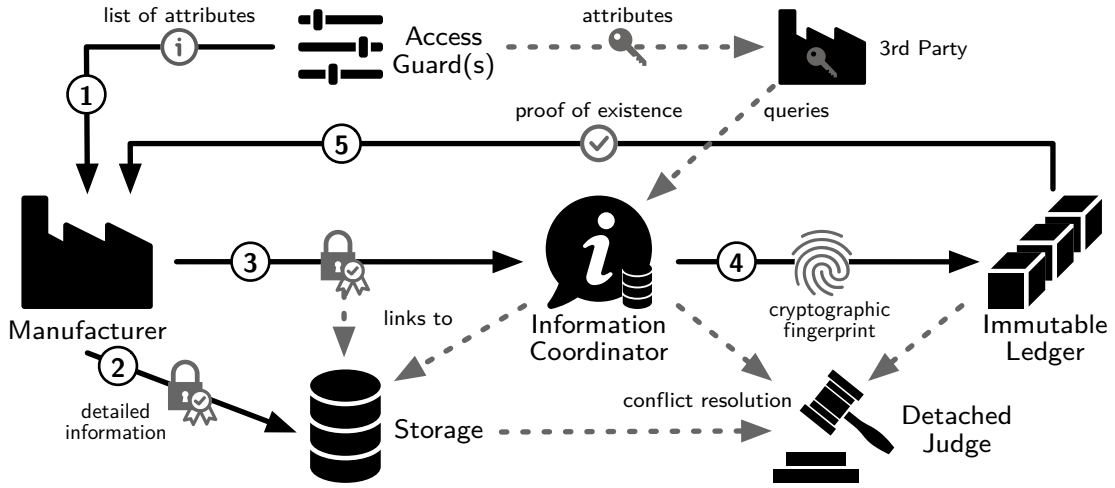



**Figure 4.4** Our design operates as follows. ① Companies retrieve a list of available attributes to configure suitable ABE policies for shared information, i.e., the attributes allow for fine-granular access control. ② They further persist shareable information at a dedicated storage before ③ provides a record that utilizes ABE and captures their interaction with a supply chain to the information coordinator. After processing (and persisting) the received record, ④ the coordinator stores a cryptographic fingerprint of the record on the immutable ledger. Finally, to have a proof of existence of the shared information, ⑤ the sharing company verifies the fingerprint on the immutable ledge with a fingerprint of the provided record. Third parties can then query the information coordinator for shared information, which they can decrypt if their attributes satisfy the configured ABE policy. A detached judge is supposed to resolve conflicts.

would be detrimental to the design's scalability (**G-P5**) and would also make the communicating entities increasingly dependent on each other (**G-P6**). To overcome both issues, PrivAccIChain relieves information-sharing manufacturers from the need of handling information flows and long-term storage themselves by securely outsourcing all information to a logically-centralized *information coordinator*. In Figure 4.4, we provide an overview of PrivAccIChain, which operates as follows. First, ①, access guards enable manufacturers to encrypt their information with ABE policies to preserve confidentiality. In particular, we rely on a hybrid encryption scheme [BPM+21, Figure 3], i.e., to improve the performance and reduce the overhead introduced by ABE, we AES-encrypt the information and only secure the AES key using ABE. For improved scalability (**G-P5**), ② PrivAccIChain optionally relies on external *storage* to outsource significant volumes of product and process information. ③ the manufacturer prepares an information record for processing by the information coordinator. This signed record contains payload information, optionally a link to a storage location, and additional details to support tracing operations. Initially, processed records only include upstream tracing information (incoming edges) that is eventually augmented with downstream tracing information (outgoing edges) once the flow of the recorded product continues.

The information coordinator only holds and processes encrypted information and conceptually acts as a single point of contact that handles all provision, retrieval, and update requests by the different supply chain actors. Additionally, a subset of PrivAccIChain participants, and optionally external contractors, jointly maintain

an *immutable ledger* where the information coordinator stores cryptographic fingerprints of information records (following ④) to facilitate their verifiability (**G-P2**) and thereby provide (long-term) accountability (**G-P1**). Subsequently, ⑤ the manufacturer verifies the correctness of the fingerprint, which also serves as proof of existence of the provided information record. Finally, a *detached judge* impartially manages disputes, e.g., for exceptional cases that require manual investigation.

Using assigned ABE attributes, information-retrieving parties can then query the information coordinator, which also implements a form of access control before sharing records. Moreover, they can only decrypt the record if they satisfy the corresponding ABE formula. In general, retrieving parties only have to interact with the information coordinator, i.e., no direct interaction is needed with the sharing manufacturer, which promises accountability (**G-P1**), verifiability (**G-P2**), scalability (**G-P5**) and autonomy (**G-P6**), especially for information flows over multiple hops, even if involved companies are defunct or non-compliant. Likewise, when tracing products, the tracing party only has to interact with the information coordinator.

In conclusion, our conceptually-centralized information coordinator enables privacy-preserving multi-hop information sharing along supply chains. While the use of ABE in PrivAccIChain ensures confidentiality, the persisted fingerprints on the immutable ledger contribute to the accountability and verifiability of processed information.

#### Components and Actors in PrivAccIChain

For an understanding of the responsibilities and duties of the different components in PrivAccIChain (cf. Figure 4.4), we now elaborate on their roles.

**Manufacturer.** For simplicity, we refer to participating supply chain actors as manufacturers, which includes both reading and writing parties. In the context of sharing and retrieving information, manufacturers only interact with the information coordinator to avoid costly point-to-point communication among multiple parties along the same supply chain. To ensure confidentiality (**G-P3**), prior to submitting (symmetrically-encrypted) information records to the information coordinator, they have to encrypt the symmetric AES key via ABE, which provides fine-granular access capabilities. Once the information coordinator confirms the reception of the record, the submitting manufacturer should check that the matching fingerprint is recorded on the immutable ledger. The persisted fingerprint then serves as proof of existence to ensure long-term verifiability (**G-P2**) and accountability (**G-P1**).

**Information Coordinator.** As a conceptually-centralized entity, the information coordinator handles all information records submitted by the manufacturers. It is responsible for (i) handling all trade and produce operations in the supply chain, (ii) submitting fingerprints of received records to the immutable ledger, and (iii) serving as an endpoint for queries by manufacturers. Despite this central component, PrivAccIChain is designed to identify misbehaving parties among both the manufacturer and the information coordinator. Through the records on the immutable ledger, manufacturers can prove that they submitted correct data to the information coordinator. Likewise, manufacturers can detect a misbehaving manufacturer

as it leads to missing or incorrect data on the ledger. We rely on this component to enable use cases, such as tracing and multi-hop information sharing, even in supply chains with dynamic and short-lived business relationships. Accordingly, its scalability (**G-P5**) is essential. Therefore, we support request batching, i.e., sharing or querying for multiple records at once. We further account for this fact by allowing for a distributed realization of the information coordinator, e.g., by partitioning specific subtrees of supply chains to different coordinators. Thus, in practice, the information coordinators are likely to be operated by an independent consortium to avoid immediate conflicts of interest and to avoid a single point of failure.

**Access Guards.** To prevent data leaks, we separate the information-handling component (*information coordinator*) from the component that enables fine-granular access control. In particular, we rely on multiple parties that serve as access guards to distribute attributes to manufacturers. This design decision prevents a single party from controlling all decryption capabilities, as each access guard only manages a subset of the attributes. Thus, this design choice is crucial to ensure privacy preservation (**G-P3**). These attributes are essential when utilizing ABE and fine-granular access policies (formulas), which are configurable for use case-specific needs. Access guards can be operated by manufacturers, but they can also be run by external parties, e.g., trade associations or governments. Overall, PrivAccIChain supports arbitrary many access guards, which are deployable for the desired level of security.

**Immutable Ledger.** As in our reliable sensing design (cf. Section 4.1.2), PrivAccIChain utilizes an immutable ledger to allow for long-term verifiability (**G-P2**) and accountability (**G-P1**). The ledger stores cryptographic *fingerprints* of all information records that have been processed by the information coordinator. Since it only stores cryptographic fingerprints, the ledger does not introduce any privacy issues. Given the ledger's tamperproofness, these fingerprints can serve as proof of existence. Moreover, incorrect or missing fingerprints allow the manufacturers to prove misconduct by the information coordinator. Furthermore, once a fingerprint has been persisted, manufacturers cannot collude with the information coordinator anymore as the fingerprint has already been persisted on the ledger (**G-P4**).

**Storage.** PrivAccIChain supports external storage, e.g., cloud storage, to outsource (encrypted) information for future use and verification. Thereby, we significantly improve its scalability (**G-P5**) over an approach that retains all information at the information coordinators. When using an external storage, the manufacturer only provides the information coordinator with an (encrypted) pointer to the information. Thus, we massively increase the flexibility of PrivAccIChain by allowing manufacturers to either store information directly with the information coordinator or with (third-party) storage providers. Certainly, this outsourcing of information might impede the overall autonomy (**G-P6**) as companies have to be complaint and further impairs the verifiability (**G-P2**) because the information coordinator does not verify linked information at the storage. To still ensure the intended reliability, we consider the manufacturer to be responsible for maintaining the required availability, i.e., if critically needed information is missing or inaccessible, we deem it the owner's fault.

**Detached Judge.** Finally, PrivAccIChain assigns a detached judge for on-demand dispute resolution for situations where information is (i) missing (or unavailable),

(ii) (purposely) incorrect, or (iii) not decryptable. Determining which entity misbehaves can be a complex and non-trivial task as it might depend on use case-specific agreements that are not captured by PrivAccIChain. As this extension is optional and for conflict resolution only, it does not impact our design's autonomy (**G-P6**).

To conclude, these six logical components are sufficient to privacy-preservingly enable accountable multi-hop information sharing along supply chains. We refer to our previous paper [BPM+21, Section 5] for a more elaborate discussion on this matter.

#### Accountable and Verifiable Information Provision, Updates, and Retrieval

To replicate the DAG structure of relationships in the supply chain network, in PrivAccIChain, we rely on doubly-linked information records, i.e., each processed record points to the preceding and succeeding trade or produce step. We refer to our previous work [BPM+21, Figure 3] for a detailed visualization of this data structure, which emphasizes the double-linked record structure.

The information sharing (provision) works as follows. First, a manufacturer submits a *produce* operation which also contains *upstream tracing (←)* information, i.e., references to a set of consumed products (records). PrivAccIChain allows both the *payload* and *tracing* information to be encrypted via AES to enforce a desired data privacy policy. The corresponding AES encryption is protected with an ABE policy that only grants access to specific manufacturers. For improved flexibility, PrivAccIChain also supports the selective encryption of nested objects with (optionally) different policies in the payload. As a next step, the manufacturer updates the *downstream tracing (→)* information of records referenced by the previously-submitted *upstream tracing* information. In contrast to *produce* records with their production details, *trade* records capture changes in ownership. *Updates* retain the history of the record to ensure the verifiability and accountability of all processed information, i.e., the information coordinator only persists the changes with respect to the previous version to minimize the storage overhead.

For (multi-hop) information retrieval, manufacturers have to only interact with the information coordinator. In particular, they query for a *produce* or *trade* record, decrypt its tracing references, and then receive the referenced records iteratively. In scenarios where tracing references are not encrypted (i.e., leaking the supply chain structure to the information coordinator), the information coordinator can independently trace products to improve performance. Furthermore, PrivAccIChain supports batch requests of multiple records at once to enable faster traversal through the DAG. As the records also include a digital signature, manufacturers can quickly verify them without involving other components or the immutable ledger. If desired or required, they can further compute a cryptographic fingerprint (at a later point) and compare it to the fingerprint that is persisted on the ledger.

### 4.1.3.5 Implementation of PrivAccIChain and Building Block Evaluation

To assess whether our design, PrivAccIChain, is suitable for real-world deployments, we created an implementation, which we briefly introduce in the following. Subse-

quently, we discuss the scalability (**G-P5**) of the utilized technical building blocks. Given the size of (modern) supply chain networks with several actors, their performance is critical when dealing with large-scale deployments. After that, in Sections 4.1.3.6 and 4.1.3.7, we evaluate PrivAccIChain with two real-world examples.

### Implementation and Experimental Setup

We implemented Python-based prototypes of PrivAccIChain's main components, i.e., for the manufacturers, the information coordinator, and the access guards. For the immutable ledger, we utilize Quorum [JPM16], an Ethereum [Woo14] fork, which supports more than 2000 TX/s (transactions per second) [BSKC18]. In contrast to Ethereum, we use the Raft [Con23] consensus algorithm, which follows the proof-of-authority concept [DAAL+17] to ensure the quick processing of transactions. We refrain from evaluating deployments that rely on external storage as (i) their performance is beyond the scope of our design, and (ii) manufacturers can select them based on their needs. We further implement our hybrid encryption scheme with the help of Charm [AGM+13], which includes implementations of AES and ABE [LW11]. We derive digital signatures, i.e., for queries and fingerprints, with the eth-account Python library [Eth18]. The information coordinator runs MongoDB [Mon09], which serves as the database backend, while Apache 2.2 [The20] handles all requests issued by the manufacturers to forward them to our implementation via ModWSGI [Dum07].

Our implementation of PrivAccIChain and an exemplary supply chain model (cf. Section 4.1.3.6) are publicly available [SrcC24b]. Due to its sensitivity, we may not share our real-world model, which is based on an electric vehicle (cf. Section 4.1.3.7).

We conduct our measurements on a server with two Intel XeonSilver 4116 CPUs, i.e., 12 cores and 24 threads each, as well as a total of 196 GB RAM. All components run on the same server to demonstrate PrivAccIChain's moderate computational footprint. In practice, they are distributed over several entities. We repeat each measurement 30 times and report 99 % confidence intervals over these runs.

### ABE Performance

In PrivAccIChain, we utilize ABE to realize fine-granular access control and confidentiality (**G-P3**). Accordingly, sufficient encryption and decryption performance are relevant for preparing, processing, and retrieving information records.

As we show in previous evaluations [BPM+21, Figure 8], the ABE performance is suitable for our setting and large-scale deployments. The (non-repetitive) ABE bootstrapping of key material, which mainly affects the distributed access guards in PrivAccIChain, is negligible. We further looked into the performance of encryption and decryption operations as part of PrivAccIChain's hybrid encryption scheme. To this end, we considered ABE policies with conjunctions and disjunctions, varied the policy length, and further included payloads of different sizes. With our ABE scheme [LW11], the number of access guards and the total number of available attributes do not influence the performance. Thus, we excluded them from our

building block evaluation. We observed that the number of used ABE attributes linearly influences the cryptographic operations, with the decryption runtime outperforming the encryption runtime. Moreover, disjunct policies are more performant than conjunct policies, which fits well to our setting with many actors. Finally, the AES-based operations on the payload are independent of the underlying ABE policy.

#### Immutable Ledger

Apart from the ABE performance, we looked at the immutable ledger [BPM+21]. In this context, we also considered design alternatives to the basic approach of issuing a dedicated transaction per fingerprint: We highlighted the potential of bundling multiple fingerprints in a single transaction, optimizing the transactions by removing redundant information, and persisting meta-fingerprints. Moreover, we discussed special concepts for the ledger, such as sharding [CDE+12]. We concluded that the ledger performance greatly depends on the underlying concrete blockchain implementation and its consensus mechanism. Accordingly, we argued that deployments can be scaled to their respective performance needs at moderate costs (e.g., slightly-increased verification overhead) without impacting the operation of PrivAccIChain.

Finally, the performance of databases is essential when assessing the scalability of PrivAccIChain. Since our design does not introduce specifically-challenging operations, we simply refer to the well-established concepts of scaling up and scaling out (the information coordinator is only conceptually a single entity in PrivAccIChain).

Hence, we conclude that the performance of the building blocks is sufficient for large-scale deployments. Therefore, we now focus on the entire design and its interplay.

### 4.1.3.6 Performance Evaluation of our Fine-Blanking Line Application

We evaluate PrivAccIChain in two different real-world settings. As our first application, we consider a fine-blanking line, including pre- and post-processing steps.

#### Supply Chain Model

As we visualize in Figure 4.5, we model a supply chain involving a fine-blanking line (cf. Section 3.2.1). During the production of a fine-blanked part, the following individual steps are involved: supplying the metal (including the processing steps of molding and rolling), operating the different parts of the fine-blanking line (i.e., coil, leveler, lubricator, and press), and several grinding, cutting, and hardening operations. Each of these steps has operating resources, tools, and a (machine) supplier, whose individual supply chains need to be taken into account. In line with an evolving industrial landscape, we further model a fine-granular separation of all processing steps (even within the fine-blanking line). Today, the separation usually comprises several pre-processing steps (e.g., molding and rolling) and subsequent assembly steps. Additionally, we assume a final product, as present in a medium-sized automobile, that combines 100 fine-blanked parts. We model the worst case

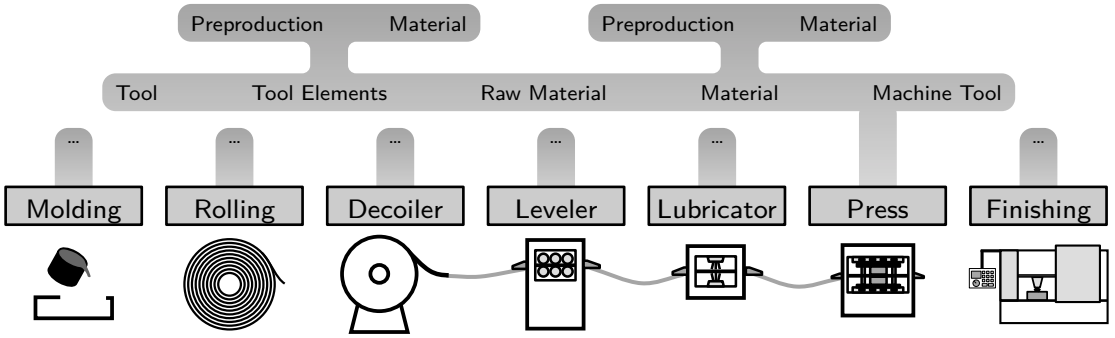



**Figure 4.5** Each processing step depends on various supply chains, e.g., to utilize tools or material. We exemplarily illustrate a subset of the (layered) dependencies for the press production step. All other production steps would showcase similar dependencies, indicated by "...".

in terms of required processing steps for PrivAccIChain by incorporating that each fine-blanked part originates from an isolated supply chain. Thus, the resulting model for our first evaluation represents a realistic (future) supply chain structure in terms of branching, depth, and total production steps.

Overall, with our *final product* that consists of 100 fine-blanked parts, we end up with a total of 410 001 nodes and 410 000 edges in the corresponding DAG, which also represents the tree-like supply chain structure in this scenario.

### Performance Measurements

Based on this supply chain model, we evaluate the real-world performance of the record creation and updating in PrivAccIChain. Subsequently, we study the use case of tracing, covering both the handling ($\rightarrow$) and sourcing ($\leftarrow$) of faults.

**Produce, Trade, and Update Records.** First, we measure the performance of individual *produce* and *trade* operations and record *updates*. We include an *encrypted produce* payload of 1 KiB and sign each operation as intended. We limit our measurement to the manufacturer and information coordinator and refer to our building block evaluation (Section 4.1.3.5) for the immutable ledger performance.

Using 20 client processes, our measurements of the internal DAG construction show that a single *produce* operation takes $85.48 \pm 0.31$ ms. Similarly, a single *trade* operation is processed in $83.92 \pm 0.14$ ms. For both operations, encrypting the payload takes more than 70 % of the time, while the information coordinator's average runtime does not exceed 4.5 ms. The time for a single *update*, i.e., providing tracking information, averages at $55.69 \pm 0.10$ ms. During our measurements, neither the information coordinator nor the underlying database operated at maximum capacity.

**Tracing.** Now, we analyze the performance of *sourcing ($\leftarrow$)* and *handling ($\rightarrow$)* a product in our supply chain model. To this end, we first performed a complete trace originating from the *final product* ($\leftarrow$), resulting in 204 800 individual product flow paths consisting of 820 001 *produce* and *trade* records. Since the runtime is driven by the cryptographic operations on the client, making the task embarrassingly parallel,

we can achieve a nearly-linear speedup by using multiple processes. For example, 30 client processes complete the tracing on encrypted data in $2553.41 \pm 19.35$ s, with still more than 80 % of the time spent on decryption. Alternatively, manufacturers could only encrypt payloads and not the tracing references to reduce the complexity, which we study in Section 4.1.3.7 along with a scenario where the tracing information of several records that are intended for the same recipient is encrypted with the same key. Regardless, performing a complete trace is a rare event as it is only of interest in exceptional cases, e.g., for in-depth investigations of plane crashes or severe malfunctions in cars. In contrast, handling faults ($\rightarrow$), i.e., tracing an initial resource to the final product, took 35 requests in $3.67 \pm 0.13$ s using a single process.

Following this first application, we observe a reasonable performance of PrivAccIChain. Given that the client (manufacturer) deals with the (parallelizable) cryptographic operations, our design's scalability is adequate, and PrivAccIChain can easily scale to large-scale settings. In the following, we validate these findings with a second application covering the supply chain of an electric vehicle.

#### 4.1.3.7 Performance Evaluation of our Electric Vehicle Application

Our second real-world application corresponds to the assembly of an urban electric vehicle (cf. Section 3.2.1) and its supply chain. As for our first evaluation, we study the performance of the record creation and updating as well as the retrieval performance. That is, we measure the runtime needed to complete upstream and downstream traces in the context of handling ($\rightarrow$) and sourcing ($\leftarrow$) of faults. Afterward, we briefly discuss the performance of PrivAccIChain.

##### Supply Chain Models

To conduct a realistic evaluation, we rely on a real-world supply chain that covers the complete product composition of an electric vehicle, which consists of more than 90 pre-assembled components, such as the body, front doors, the rear axle, the battery, and the engine. We refer to our previous paper [BPM+21, Figures 6 and 7] for a visualization of the entire supply chain and the vehicle's composition. For our evaluation, we consider two models with different complexity. In our *base model*, we represent each component of the vehicle as an individual product of the product flow DAG, which corresponds to the creation of a *produce* record. Hereby, we deviate from the real structure and assume that each production step is executed by a unique company, such that we include ownership transfers represented by *trade* records between each production step. Therefore, our evaluation again covers the worst case in terms of the required processing steps for PrivAccIChain.

In contrast to the fine-blanking application, the resulting DAG structure is broader and features a more irregular branching behavior. Moreover, it comprises only 10 instead of 17 levels. Despite this difference, we assess this model as realistic as well because the source nodes on long paths correspond to basic components, such as screws, nuts, or rivets—the so-called "c-parts". Overall, for this base model, we obtain a DAG with 2222 nodes and 2221 edges. However, in the base model, we also

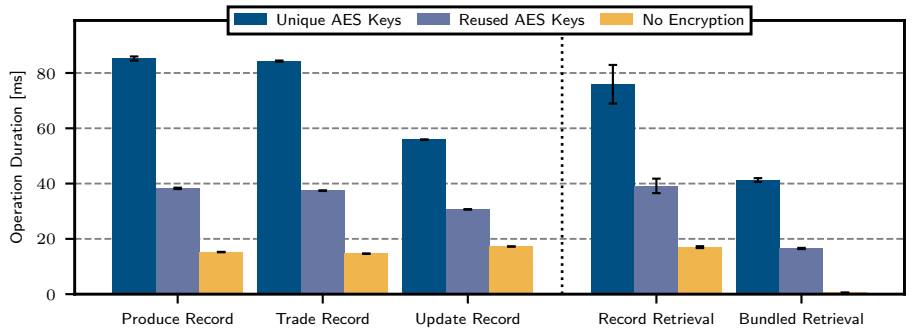



**Figure 4.6** Operation runtimes of our electric vehicle application for record provision, updates, and retrieval for three encryption settings. The cryptographic operations performed by the manufacturers dominate the processing of each operation. Retrieving multiple records in a bundle provides a tunable trade-off between record retrieval latency and record throughput.

identified source nodes with a path length of three or four that represent different pre-assembled components, such as electronic control units produced by external suppliers. Thus, to further decompose them, we derived a second, *extended model* to also cover the production steps of these pre-assembled components. To this end, we append seven full ternary DAGs of depth three to each source on these two levels. This addition results in 280 additional nodes for each source node. The resulting extended model consists of 133 822 nodes and 133 821 edges overall. Consequently, it is comparable in scale to our previous evaluation of a fine-blanking line.

### Record Creation, Updates, and Retrieval

First, we again look at the processing of new or updated *produce* records with a payload size of 1 KiB. In contrast to our initial evaluation, we now consider three settings to study the trade-off between confidentiality and performance: (i) utilizing no encryption for the record's payload and the tracing references (as a baseline), (ii) applying record encryption for the payload and the tracing references while reusing AES keys along the supply chain, which allows manufacturers to cache the resulting ABE-encrypted keys for increased performance, and (iii) applying encryption with unique AES keys for each record and record field, i.e., we encrypt the tracing references and payload with unique AES keys per record. Consequently, we compare an (insecure) baseline scenario with an optimized but realistic encryption configuration and a worst-case scenario. To increase the load on the information coordinator, we parallelize the creation and updating operations by using 20 processes. Next, we discuss the corresponding results, which we also summarize in Figure 4.6.

**Record Creation and Updates.** Our measurements indicate that the performance of *produce* and *trade* records is comparable across all three settings. The higher deviation for *produce* records follows from varying branching conditions in the evaluated DAG that affect the included tracing references. When reusing AES keys, we achieve a speedup of more than 50 % in comparison to our worst-case setting while still encrypting the records. For both settings that involve encryption, the information coordinator takes less than 5 ms on average to process requests, which underlines

that the primary load is on the manufacturers, indicating the absence of scalability issues (cf. **G-P5**). This observation is further backed by the record creation times of unencrypted records, which take around 15 ms on average. We further measure the time required to update records. Overall, we measure an average update time of $55.88 \pm 0.12$ ms in our worst-case setting (more than 800 000 updated records). With reused AES keys, an update takes only $30.67 \pm 0.07$ ms, which is further reduced to $17.24 \pm 0.05$ ms in our baseline setting. Thus, the handling of updates is slightly slower than the creation of records because the information coordinator has to load and process the previous state to maintain a consistent version history.

The total construction time of the complete DAG is irrelevant to assess the performance because, in real deployments, the respective operations would not be executed without delay, nor would they be triggered by a single entity as each actor interacts with the information coordinator individually after processing (and shipping) products. Hence, only the performance of the information coordinator is of interest.

**Record Retrieval.** In addition to providing information, manufacturers also want to retrieve records to access shared information. Instead of issuing a single query for each record, manufacturers batch their requests in a single query. While this design increases the complexity of this query, it reduces the overhead in terms of request signatures and round trip times, such that we expect a significant performance benefit for the effective retrieval time per record. In Figure 4.6, we compare the retrieval times of single record queries and batched queries that request 100 requested records per query for all encryption settings. Requesting a single record takes $75.93 \pm 6.98$ with unique AES keys, $39.16 \pm 2.62$ with reused AES keys, and $17.01 \pm 0.34$ for unencrypted records. Batching significantly improves the effective retrieval time per queried record: For batches of 100 records, we measure per-record averages of $41.27 \pm 0.69$, $16.52 \pm 0.22$, and $0.53 \pm 0.06$, respectively, for the three encryption settings. Especially unencrypted records profit from bundling. Since our measured retrieval times do not exceed 85 ms, we generally assess PrivAccIChain's performance in this regard as performant and real-world applicable.

### Tracing

As for our first application, we now look at the use case of tracing to further study the performance in supply chains with multi-hop information retrieval. A complete trace (upstream $\leftarrow$) that originates from the *final product* consists of 2222 *produce* and 2221 *trade* records for our base model. For the extended model, the corresponding full trace to source faults consists of 133 822 *produce* and 133 821 *trade* records. In the other direction (downstream $\rightarrow$), we identify the longest path with 10 levels in our DAG. We consider this worst-case situation when measuring the performance of handling faults, i.e., we provide an upper bound.

**Sourcing Faults ($\leftarrow$).** We parallelize the tracing with 30 client processes for our settings with encryption to parallelize the decryption of records. For our unencrypted baseline setting, we trace using 10 threads and a single process because the resulting synchronization overhead of several processes degrades the overall performance. In

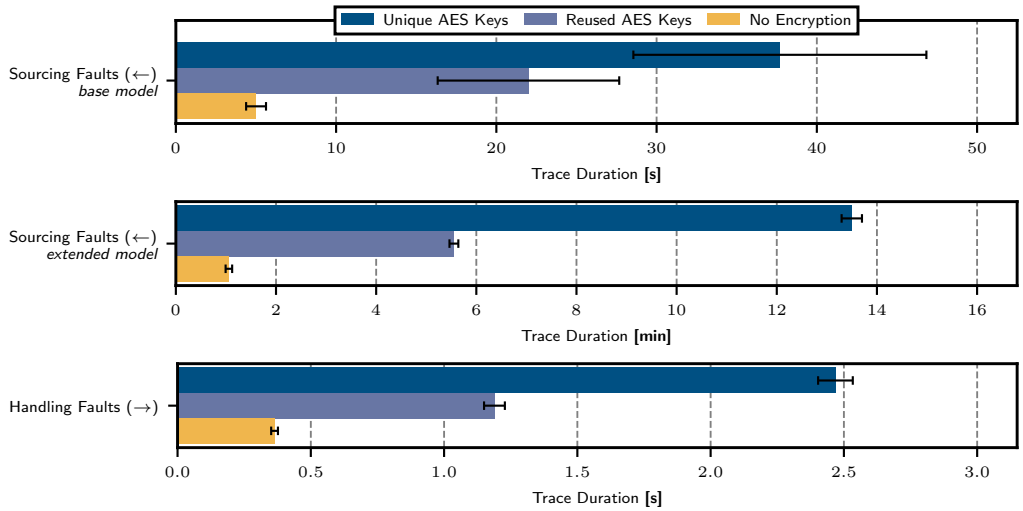


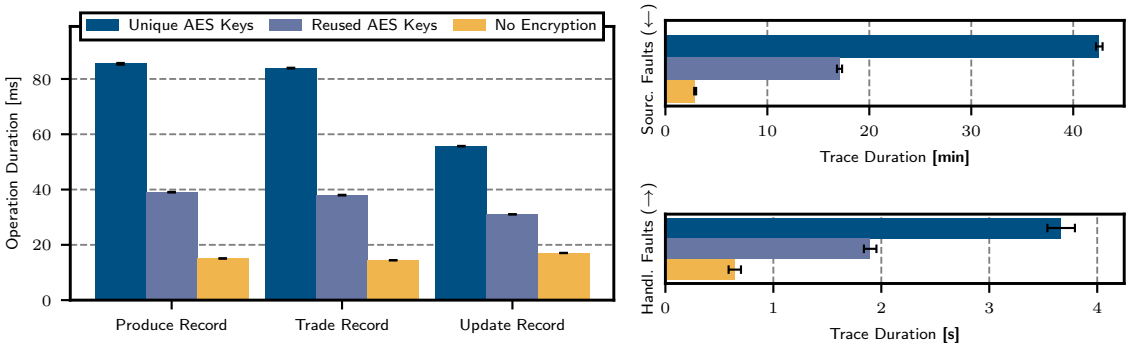

**Figure 4.7** Similar to record provision and retrieval, the tracing duration is dominated by the required decryption operations at the manufacturer. As the longest path is identical for both models of our electric vehicle application, the worst-case runtime to handle faults is identical.

all settings, we further batch up to 500 records in a single request. For both models, we visualize the sourcing of faults in Figure 4.7; however, we illustrate them with different scales for improved readability. With the base model, we measure $37.69 \pm 9.14\,\text{s}$ for uniquely-encrypted records, $22.00 \pm 5.66\,\text{s}$ for records encrypted with reused AES keys, and $5.01 \pm 0.62\,\text{s}$ for unencrypted records. For the extended model, a full trace takes $809.89 \pm 12.19\,\text{s}$ with unique AES keys, $333.15 \pm 5.30\,\text{s}$ for reused AES keys, and $63.54 \pm 3.91\,\text{s}$ for unencrypted tracing references.

As full traces in supply chains are rare events, e.g., in case of accidents or food poisoning, the measured performance is reasonable even in the worst-case setting with unique encryption keys for each record. Since the retrieval of all records from the information coordinator is concluded after a minute or less (most time is needed for cryptographic operations at the manufacturer), increasing the manufacturer's computational resources is sufficient to speed up the tracing. Parallelizing the sourcing of faults further increases the performance as subgraphs are traversable independently, allowing for distributed and horizontally-scalable tracing implementations.

**Handling Faults (→).** In the context of handling faults, we trace the longest path of the DAG as an upper bound. As we illustrate in Figure 4.7, the corresponding tracing runtime is negligible as we measure only $2.47 \pm 0.07\,\text{s}$ with unique AES keys, $1.19 \pm 0.04\,\text{s}$ with reused AES keys, and $0.36 \pm 0.01\,\text{s}$ with unencrypted tracing references. When handling faults, every record query can only request a single record. Hence, we cannot benefit from batching, regardless of the respective DAG.

While tracing is somewhat constrained by the manufacturer-driven querying (and the need to decrypt tracing references), our design ensures that the information coordinator remains oblivious of the tracing references (when using encryption). Although we consider these references to be sensitive, in scenarios with established trust, coordinator-driven tracing constitutes a reasonable alternative as it allows for significant performance improvements. The information coordinator can simply




**(a)** Runtimes for record provisions and updates

**(b)** Tracing durations

**Figure 4.8** The runtimes (including tracing) for the three encryption settings of our fine-blanking line application are comparable to the performance of our electric vehicle application.

perform the complete trace independently, as no time-consuming round-trips to the querying manufacturer are needed to decrypt encrypted tracing references.

### Evaluation Discussion

For this evaluation, we considered two models of different scale with different encryption settings to provide insights into the performance implications of strict confidentiality needs. Despite the significantly-different impact of the encryption settings, we assess the performance for these operations as appropriate, as all operations take less than 100 ms on average, with the majority of computations and runtime spent by querying clients. A comparison to our application of a fine blanking production line confirms these findings across all encryption settings (Figure 4.8). The overhead of applying cryptographic operations is mostly limited to the manufacturers. Therefore, we conclude that PrivAccIChain scales well with an increasing number of operations and manufacturers. Given that the use and level of encryption are configurable on a per-record basis, PrivAccIChain achieves a tunable trade-off between privacy and performance while supporting multi-hop information sharing. As individual operations are in the orders of milliseconds, PrivAccIChain is well-suited for real-world use. In real-world deployments, production and shipping steps realistically take hours, days, or even weeks, in contrast to our evaluation, where we trigger them instantaneously when processing all produce and trade records that belong to our DAG, which represents the entire supply chain (network).

We further ascertain that the underlying supply chain structure, i.e., the branching behavior and depth of the DAG, does not have a major impact on the tracing performance at larger scales. Compared to our fine blanking application that considered longer paths, yet sparser branching behavior, we observe linear scaling regarding the tracing durations. Hence, we conclude that for supply chains with $n$ nodes, which exceeds the number of utilized processes $p$ by several magnitudes (i.e., $n \gg p$), only $n$ has a significant impact on the tracing duration. Finally, we refer to our previous paper [BPM+21] for a comparison of PrivAccIChain to ProductChain [MKJ18], another state-of-the-art approach in the area. There, we have shown that our results align with other designs for tracing in supply chains. However, in contrast to our

work, ProductChain assumes trustworthy manufacturers, and therefore, it is not applicable to our considered scenario with low-trust supply chain environments.

After the (computational) performance, we now discuss PrivAccIChain's security.

#### 4.1.3.8 Discussing the Security of PrivAccIChain

In addition to the performance evaluation, we further have to discuss PrivAccIChain's security to holistically assess its real-world feasibility. We utilize several established cryptographic and technological concepts, such as attribute-based encryption, symmetric encryption, or blockchain technology. Thus, we rely on the individual security of these concepts as well as their respective implementations, namely AES [DBN+01], the utilized CP-ABE scheme [LW11], and Quorum [JPM16], as well as all utilized libraries, e.g., the Charm cryptography framework [AGM+13]. We further exclude external factors, such as long-term power outages, as out of scope. Accordingly, we now focus on different attack vectors against our design while considering malicious-but-cautious entities (cf. Section 2.1.2.1).

##### Attack Vector Analysis

In the following, we consider and discuss various attacks with differing likelihood and severity. We order our presentation according to their decreasing severity.

**Entity Collusion.** Entity collusion refers to attempts of multiple entities within our architecture to gain access to information in violation of established access policies. Since information is stored in (encrypted) records on the side of the information coordinator, the information coordinator has to be involved in the collusion to provide the records. Encrypted records are only decryptable if the required ABE attributes are available. Based on this requirement, we identify two potential scenarios.

First, a manufacturer satisfies the ABE policy. Then, it legitimately has access to all attributes that are required for decryption, i.e., no illegitimate access occurs. Due to the collusion resistance of the underlying ABE scheme, joining attributes of different manufacturers will not lead to increased decryption capabilities, i.e., the information remains encrypted [LW11]. Second, the issuing access guards could collude to obtain additional decryption capabilities. In contrast to a collusion of multiple manufacturers, PrivAccIChain cannot prevent this type of attack. However, PrivAccIChain allows for a tunable collusion resistance because the issuing of ABE attributes can be distributed over multiple access guards. Moreover, appropriate ABE policy design promises to further reduce the threat of this attack vector, i.e., attributes that satisfy a policy should be distributed over multiple access guards.

**Intentional Data Distribution.** Besides the collusion of multiple entities, records could be shared with unauthorized parties on purpose. If encrypted records are distributed, no information is leaked as sensitive information simply remains encrypted. In contrast, if a manufacturer with legitimate access decrypts the information, this information is not processed and secured by our design anymore. Thus, PrivAccIChain cannot prevent its (unauthorized) distribution. Here, we recommend using access logs at the information coordinator to identify misbehaving parties.

**Record Tampering.** Manufacturers have to trust the information coordinator to store their records persistently. Thus, the availability and correctness of these records depend on the information coordinator's compliance. Hence, unintentional data loss as well as intentional manipulation or deletion of records are also attack vectors. Backup and replication strategies promise to mitigate the risk of (unintentional) data loss. Similarly, the fingerprints on the immutable ledger ensure that manufacturers can prove intentional manipulations and deletion of records. Thus, we conclude that PrivAccIChain supports sufficient mechanisms to counter record tampering.

**Illegitimate Behavior.** Our design assumes that the majority of entities behaves as intended, i.e., they are not colluding to manipulate, delete, or distribute data. Consequently, majoritarian illegitimate behavior can impact PrivAccIChain.

By distributing the control over ABE attributes to multiple access guards, encrypting records, storing proofs of existence (fingerprints) on the immutable ledger, and allowing multiple instances of the information coordinator, PrivAccIChain provides several countermeasures against entity misbehavior and collusion. The level of resilience against illegitimate behavior is tunable at the cost of increased complexity and decreased performance. However, PrivAccIChain is prone to access guards colluding, for example, with the information coordinator, i.e., where a majority of central components colludes or misbehaves. In such a case, the confidentiality of records is not technically guaranteed anymore. While this attack vector could pose a significant threat, we rate its likelihood as very low because manufacturers will simply cease to rely on the respective (compromised) instance of PrivAccIChain.

**Information Fraud.** While PrivAccIChain protects against record manipulation and deletion, it cannot guarantee the correctness of provided information. Thus, manufacturers could insert fraudulent information, e.g., manipulated information or random data. To account for this attack, PrivAccIChain features a detached judge for on-demand conflict resolution, for example, to express financial penalties. Our reliable sensing (cf. Section 4.1.2) or integrated reputation systems (cf. TrustChain [MDKJ19]) could provide additional resilience in this regard.

**Request Pattern Analyses.** In addition to illegitimate and direct access to information, attackers could be interested in business relationships and frequent interactions, for example, when eavesdropping the interactions with the information coordinator. Although PrivAccIChain allows for record encryption, request and communication patterns (e.g., request frequencies and volume) are still eavesdroppable. Moreover, the information coordinator has access to these patterns by design. To counter such passive attacks, manufacturers can manage multiple accounts to appear as if different manufacturers are interacting with PrivAccIChain. Alternatively, obfuscation strategies, e.g., the use of dummy records or mix networks, could be appropriate countermeasures that do not excessively burden the information coordinator. Thus, corresponding mitigation strategies are conceptually available if needed.

**Key Leakage.** PrivAccIChain relies on AES and ABE to ensure confidentiality, i.e., records are not encrypted individually for each recipient. In addition to reducing the overhead, this design also ensures that single manufacturers cannot be excluded from access to information (if they satisfy the embedded ABE policy). Given that ABE

attributes are bound to individual manufacturers, a collusion of multiple entities will not lead to elevated access either [LW11]. Thus, only a single leak by one party with all required ABE attributes would provide decryption capabilities. However, the information coordinator does not return any records to unauthorized parties. Thus, even in this case, the implications of key leakage attacks are limited.

Similarly, compromised access guards are only a minor threat because ABE policies should be designed in a way that they require attributes by $n$ different access guards. Consequently, no single party is responsible for all attributes. Additionally, all entities authenticate themselves using certificates. If the corresponding key material is compromised, an attacker could impersonate another entity. However, only with simultaneous access to the relevant ABE attributes, the attacker would be able to decrypt retrieved records. With the help of request origin determination or request logs, such an attack should be detectable to allow for certificate revocation.

Altogether, key leakage represents a significant attack vector with only limited severity due to our multi-layered and decentralized access control scheme. The implementation of advanced mechanisms, such as time-interval attributes [LYZL18], could further reduce the implications of key leakage attacks.

**Denial-of-Service Attacks.** Apart from specialized attacks, denial-of-service attacks could (also) impair the availability of PrivAccIChain. We consider corresponding mitigation strategies as out of scope and refer to appropriate countermeasures in both past [FS00,MR04] and current research [RKKV17]. Resource accounting as well as pattern and anomaly detection represent preventive and reactive methods. Due to our distributed design and the processing of usually time-uncritical information, corresponding attacks only pose a minor threat to PrivAccIChain.

**Data Leakage.** Similar to the intentional distribution, records could be revealed following an unintended distribution or as a result of compromised entities. Although we assess the likelihood of external attacks as higher than for intentional distribution, external attacks usually do not reveal the required decryption keys. Hence, attackers can only gain limited access to valuable information (if decrypted at all).

### Conclusion of our Security Discussion

Concerning the security guarantees of PrivAccIChain, we presented nine groups of attack vectors and analyzed their severity as well as their respective likelihoods, as we summarize in Table 4.2. We have further outlined several implemented as well as optional countermeasures. Overall, we consider information fraud as the most likely attack with potentially far-reaching implications since manufacturers could expect competitive advantages by providing fraudulent information. As part of our information-processing pipeline, we intend to mitigate this threat by establishing the notion and concept of reliable, i.e., E2E-secured, sensing (cf. Section 4.1.2). Due to the tunable resilience against entity collusion in PrivAccIChain by utilizing multiple access guards, we assess the remaining attack vectors as unlikely and rather uncritical. Thus, we conclude that it provides appropriate security capabilities and guarantees as well as a tunable trade-off between accountability (transparency) and confidentiality (privacy) for information sharing along supply chains.

| Attack | Severity | Likelihood | Countermeasures |
|---|---|---|---|
| Entity Collusion | ● | ◑ | Multiple Information Coordinator Instances, Multiple Access Guards, *External Supervision* |
| Intentional Data Distribution | ● | ○ | Request Logging, *Financial Penalties* |
| Record Tampering | ● | ○ | Backup Strategies, Multiple Information Coordinator Instances, Fingerprints, Proof of Existence |
| Illegitimate Behavior | ● | ○ | Multiple Access Guards, Multiple Information Coordinator Instances |
| Information Fraud | ◕ | ◕ | Record Fingerprints, Financial Penalties, *Reputation System*, *IoT Sensor Data* |
| Request Pattern Analyses | ◑ | ◔ | Multiple Accounts, *Active Obfuscation* |
| Key Leakage | ◑ | ○ | Layered Access Control, Time-Interval Attributes |
| Denial-of-Service Attacks | ◔ | ○ | *Preemptive and Reactive Avoidance* |
| Data Leakage | ○ | ◑ | Data Record Payload Encryption |

**Table 4.2** Summary of potential attack vectors against PrivAccIChain. For each attack, we state the estimated likelihood and severity in case of a successful attack as well as all implemented and optional (highlighted in italics) countermeasures. We rate the likelihood and severity from low ○, over medium-low ◔, medium ◑, and medium-high ◕, to high ●.

#### 4.1.3.9 Conclusion

This section concludes the second part of our processing pipeline for reliable information along supply chains. We have discussed PrivAccIChain, a design that establishes accountable-yet-confidential information flows along supply chains. PrivAccIChain particularly considers environments with flexible and highly-dynamic business relationships through its use of ABE policies. Moreover, the (encrypted) tracing references in the information records account for multi-hop information sharing and further allow for efficient traversal of the supply chain. The traversal is especially beneficial for the use case of tracing, both downstream ($\rightarrow$) and upstream ($\leftarrow$). Based on our evaluation of two real-world applications, we conclude that PrivAccIChain's performance is satisfactory for large-scale deployments. With the presented flexibility of PrivAccIChain, i.e., its customizability regarding accountability, verifiability, and confidentiality (privacy) at the expense of computational complexity and storage requirements, we provide a tunable and powerful design. Finally, corresponding deployments could even support new businesses in bootstrapping their operations as PrivAccIChain provides desired accountability for them by default.

### 4.1.4 Takeaways and Future Research

In the following, we briefly conclude the presentation of our first contribution, which addresses the need for reliable information, its processing, and sharing along supply chains. In particular, we have demonstrated two (independent) parts that jointly tackle this challenge, from the "edge", where information is being sensed, to the long-term use, where information is being persisted and shared (cf. Figure 4.2). We wrap up by discussing the suitability of our selected building blocks in Section 4.1.4.1. Subsequently, in Section 4.1.4.2, we briefly highlight the universality of our presented

pipeline for applications that exceed the processing of information along supply chains. Finally, in Section 4.1.4.3, we discuss potential next steps and future work.

#### 4.1.4.1 Suitability of Selected Technical Building Blocks

Concerning the reliable sensing part, we conclude that our selected building blocks are largely without promising alternatives: Confidential computing and the idea of trusted sensors are exceptionally well-suited to improve the authenticity and reliability of sensed information as confidential computing has been proposed to protect sensitive information during processing (in untrusted environments). Our concept further relies on an immutable ledger to ensure long-term information integrity and consistency. While we have evaluated the performance of a blockchain, i.e., Quorum, in practice, any type of tamperproof storage could be utilized, such as approaches that realize database functionality inside of TEEs [RAA+19, PIZ+20]. Alternatively, all stakeholders could also agree to rely on a trusted third party instead. These changes to the long-term storage do not impact the security guarantees of our E2E-secured sensing, as sensed information only leaves the protected computing environment after it has been attested and signed by the trusted server.

When looking at privacy-preserving information sharing, we are confident to have selected tunable and performant technical building blocks, i.e., they allow for deployments of PrivAccIChain that are configurable according to use case-specific needs. As we have outlined in our previous work [BPM+21], several approaches rely on blockchain technology to improve the exchange of information in environments that deal with supply chains. However, many of these works utilize a blockchain as a central entity and require all information to be stored on-chain. To address this bottleneck, in PrivAccIChain, the immutable ledger (again, it can be any type of tamperproof storage) is only a supporting component. Thereby, we ensure the design's scalability in large-scale and highly-dynamic environments. The application of ABE as a generalization of traditional public-key cryptography further accounts for opaque information-sharing scenarios that cover both varying business partners and indirect business relationships. Without this modern concept, ensuring confidentiality would be cumbersome and computationally costly as many recipients (especially downstream) are not necessarily known when provisioning information records. The uncertainty of recipients further significantly impairs the application of other building blocks, such as data usage control or secret sharing.

When combining this discussion with our evaluation, we conclude that our selected building blocks are well-suited to explore the evolution of information sharing along supply chains in the IIoT. Moving on, we look at applications beyond this scope.

#### 4.1.4.2 Universality of our Processing Pipeline for Reliable Information

While we envision deployments featuring our full processing pipeline (from trusted sensor to PrivAccIChain) for the most extensive benefits, the individual parts are also ready for individual use. Thus, supply chains can apply our design(s) as needed.

While we motivate our reliable E2E-secured sensing with supply chains, our work is also relevant for other applications. In particular, we can easily translate the foundation of our work to settings where mutually-distrustful parties sense information in (remote) environments, e.g., to reliably handle shared inventory management, rental or parking services, digitized construction sites, and smart manufacturing. Especially the latter demands accurate processing of usage and state information due to the emergence of digital factories (cf. Section 1.1), where manufacturing equipment and raw material are shared among companies. For example, accurate monitoring of tool wear can be crucial to avoid significant damages (and, in turn, costs) to production lines. Consequently, we argue that our work can also impact applications that exceed the discussed tracking and monitoring in supply chains.

Likewise, we have designed PrivAccIChain with supply chains in mind. Even though we specifically focused on the use case of tracing, PrivAccIChain is beneficial for various types of information sharing along supply chains. From a security perspective, companies could even rely on carefully-designed ABE policies to realize information sharing across supply chains. Thus, we conclude that PrivAccIChain is a powerful concept to facilitate vastly different types of privacy-preserving information sharing.

Finally, we would like to point out that our processing pipeline is only able to provide a solid and secure foundation for reliable information based on technical building blocks. Thus, to resolve situations with deliberate misbehavior, technical failures, or legal questions, our designs still inherently require human decision-makers in the loop to make judgments and decisions based on the output of the technical domain.

#### 4.1.4.3 Future Work and Next Steps

To further improve our contribution and its long-term implications, we discuss potential next research directions and steps based on three conceptual groups.

First, a number of minor improvements, closely related to our work, are likely to ease the deployability of our processing pipeline. First, large-scale deployment of our reliable sensing design could help to convince stakeholders that relying on confidential computing for their sensing activities entails significant benefits. Besides, to consider all types of logical actors (cf. Section 4.1.1.1), exploring in detail whether appropriately-designed ABE policies in PrivAccIChain address all needs for (governmental) oversight or public verifiability, e.g., as required and desired in food chains, would be supportive as well. Moreover, guiding stakeholders in setting up our processing pipeline for real-world use in, for example, key management, remote attestation, ABE policy design, or other parameters, such as rate-limiting parameters when interacting with the information coordinator, could further boost their willingness to deploy and rely on it. Finally, when having their confidentiality concerns in mind, extending PrivAccIChain with a time-based ABE scheme could satisfy their privacy needs in settings where (long-term) transparency is considered to be less important.

Second, we identify two conceptual improvements that potentially significantly impact our contribution. On the one hand, future work could combine the concepts of PrivAccIChain and CCChain [WMP+22] to reduce the overhead of enabling long-term information sharing by only publishing proofs. When evolving our processing

pipeline in such a way, research should consider all stakeholders, including businesses, consumers, and governments. On the other hand, the development of reputation and rating systems that source reliable (and verifiable) information from our processing pipeline could reshape business relationships in industry [BPT+23]. For example, with reliable ratings, which would be based on technical guarantees, the dynamic selection of suppliers could gain further traction in the evolved industrial landscape.

Third, our work would further benefit from progress in orthogonal research directions. In particular, as pointed out by related work (cf. Section 4.1.1.2), we expect that significant effort is required in the future to develop reliable solutions for scenarios where an embedding or attachment of trust anchors to a product is impossible or infeasible (e.g., for cost or practicality reasons). Our work depends on such concepts to match physical products with their digital information in our processing pipeline. Additionally, given the focus of this dissertation on the information security perspective, we have neglected the economic implications of our contribution. Primarily, PrivAccIChain is of interest here because it could enable attribute commercialization, for example, to develop new business models: Companies could sell access rights to specific (sensitive) information when designing ABE policies in a skilled manner, effectively establishing data markets. Finally, we look forward to seeing real-world deployments that pick up our ideas despite the multitude of adoption challenges. For an overview of these challenges, we refer to Gonczol et al. [GKHD20], who compiled an extensive list based on related work of both academia and industry.

This subsection concludes the presentation of our first contribution, which focused on existing business relationships. Specifically, we have exemplified a powerful processing pipeline to implement secure collaborations, i.e., reliable information sensing and sharing, along supply chains, even in opaque supply chain networks. We now shift to our second contribution, which explicitly considers the confidentiality needs in settings that involve unknown, most likely untrusted entities along supply chains.

## 4.2 Finding New Suppliers with Privacy-Preserving Purchase Inquiries

In our second contribution, we look at secure collaborations that address the challenge of finding new suppliers, as prevalent in procurement processes. Effectively, this practice frequently depends on establishing new business relationships along the supply chain among mutually-distrusting companies. In this setting, the evolution toward dynamic business relationships (cf. Section 1.1) is likely to only succeed if technical approaches securely support companies in the establishment of new (trust) relations by providing reliable guarantees. Ideally, corresponding approaches barely introduce any overhead for the involved parties. In the following, we specifically focus on the potential of sourcing goods across the complete industrial landscape, i.e., to also consider suppliers without an established relationship, as the fear of disclosing sensitive information during procurement pushes companies to their fixed networks of suppliers [DGP03]. To the best of our knowledge, we are the first to tackle the confidentiality concerns of this essential task by introducing technical guarantees.

First, in Section 4.2.1, we describe these privacy challenges in detail. Afterward, in Section 4.2.2, we detail our designs for two-way privacy in the early stages of procurement. We refer to this step of procurement as "purchase inquiry". Finally, in Section 4.2.3, we conclude the presentation of our second contribution.

## 4.2.1 Challenges in Bootstrapping Collaborations

Companies and production processes in an evolved industrial landscape are also driven by customer requests or small-batch production (cf. Section 1.1). A necessity for any production is the availability of parts and components. Likewise, swiftly reacting to new circumstances and customer requests can only succeed if companies have sufficiently large networks of suppliers [SGM06, SOM14]. Given the diversity of requests in small-batch production, companies likely have to flexibly source parts from different suppliers. Therefore, they have to repeatedly discover suitable business partners, a process that many stakeholders consider to be very privacy-invasive.

As a foundation for our second contribution, in Section 4.2.1.1, we thus introduce purchase inquiries and refer to the privacy needs that we have already outlined in Section 3.2.2. This introduction is essential to understand the confidentiality needs of involved stakeholders. Subsequently, based on this information, we derive general design goals (Section 4.2.1.2) that promise to mitigate this situation. Afterward, in Section 4.2.1.3, we discuss related work and highlight that related work has mostly overlooked this research direction so far. Based on this foundation, we then detail our designs that tackle this overlooked yet crucial aspect of procurement.

### 4.2.1.1 Purchase Inquiries

With our focus on the IIoT in this dissertation, we particularly discuss procurement and purchase inquiries from a perspective of business-to-business markets. As part of the procurement process [NS91, WR17], interested companies, so-called *buyers*, contact potential suppliers (*sellers*) and inquire about specific components, parts, or products [CRZ20]. While we refer to this step as a *purchase inquiry*, it is also known as "request for quotation" [Elg12, CRZ20]. The primary goal of this step is to find a supplier who can satisfy the requested order. Additionally, relevant supplier evaluation criteria include, e.g., timeliness, quality, sustainability, or price [CRZ20, PMK+21]. Most notably, a purchase inquiry is only a single step in the process of identifying, managing, and integrating suppliers. For an in-depth introduction to the prevalent steps of today's well-established procurement processes in business-to-business markets, we refer to our summary in previous work [PDF+23, Appendix A].

During procurement, a buyer commonly has to consider (and contact) various suppliers, as we illustrate in Figure 4.9. For instance, she is interested in a specific product, e.g., a robotic arm, and contacts her network of suppliers (i.e., suppliers the buyer has worked with in the past). To this end, she has to provide them with detailed specifications about the robotic arm, such as rotation angle, lifting capacity, and supported communication protocols. After their replies, which indicate whether

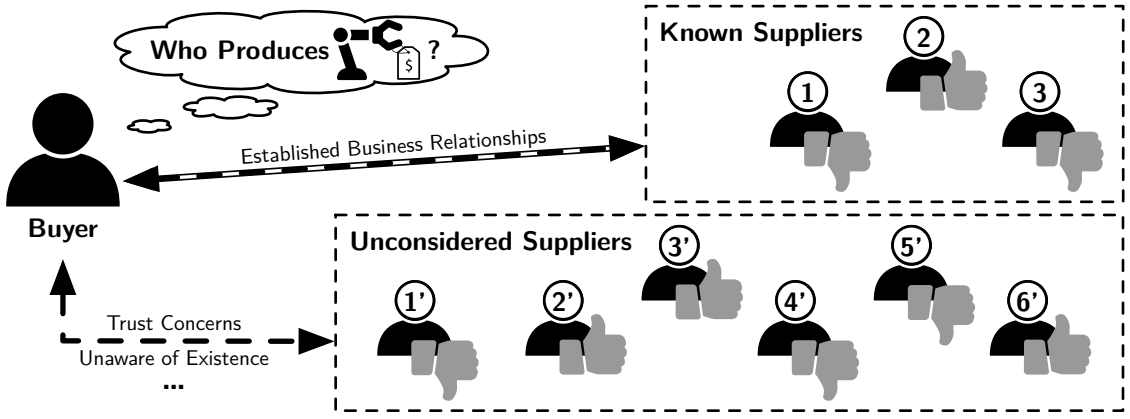



**Figure 4.9** Traditionally, buyers tend to reuse their existing business networks when looking for suitable suppliers. Due to privacy concerns (the need to share information upfront), other, potentially superior suppliers are excluded, which diminishes the result and value of the inquiry.

they are able to deliver the requested product, the buyer might be able to select a suitable supplier. However, due to the focus on (trusted) suppliers, several other potentially superior suppliers are left out. Related work confirms that finding suitable suppliers is a major issue [HP01]. Centralized platforms promise to mitigate this situation by simplifying this matching [HP01]. However, they learn all details, which constitutes a significant breach of privacy. Even without such platforms, buyers have privacy concerns, i.e., they want to avoid sharing sensitive information with companies without a previous relationship [Age01, AN07, NN17]. Furthermore, they might not even be aware of their existence or supported capabilities. Thereby, when making such decisions, buyers currently have a limited view of all available options.

Likewise, potential sellers are also dissatisfied with sharing sensitive information upfront. As their counterpart, they have to openly share details on their capabilities, their delivery times, and price expectations even if no trust relationship has been established. This knowledge could potentially benefit their competitors to outmatch their offers. Therefore, privacy in purchase inquiries is a two-way street [AN07] and should be treated accordingly. Otherwise, the risk-free establishment of relationships with the goal of selecting the most suitable supplier(s) will not succeed.

Zeng et al. [ZWD+12] outline the different types of (sensitive) information in detail and hence confirm the need for corresponding approaches. To overcome the issue of neglecting many suitable business partners, privacy preservation during procurement is thus a crucial aspect. A corresponding solution should mitigate this issue, address any concerns, and improve today's established "buyer-seller" matchmaking. Apart from an improved matching, further benefits could follow from the handling of specialized products [FR01] and the ability to swiftly react to customer requests.

In Section 3.2.2, as part of our introduction of procurement in the context of machine tool suppliers, we have already outlined the most important privacy aspects. To summarize, the process of requesting offers is characterized by the distrust of the involved parties. To receive a useful offer, both companies have to exchange sensitive information, which they prefer not to disclose, especially when dealing with untrusted parties. Consequently, we argue that technical means are needed to suffi-

ciently address the stakeholders' confidentiality needs. Ultimately, such approaches would contribute to fairer and industry-wide competition while allowing companies to also focus on less tangible goals, such as sustainability (cf. Figure 1.1). Thus, overall, the benefits exceed far beyond the directly-involved buyers and sellers.

#### 4.2.1.2 Design Goals for Improved Privacy during Procurement

Based on our considered scenario, we now derive a set of five distinct, general design goals, which must be addressed by any approach that improves the privacy of purchase inquiries. We argue that they are independent of the domains in which the inquiries take place. These goals summarize the needs of the individual parties (**G-I1** and **G-I2**) as well as universal conceptual design goals (**G-I3**, **G-I4**, and **G-I5**).

**G-I1: Buyer Privacy.** Companies are interested in keeping sensitive information on their business practices and orders private [AN07, NN17], especially in light of untrusted third parties. Consequently, they are only willing to reveal this information to their suppliers, i.e., in our scenario, the deliberately-selected seller. They want to avoid sharing anything upfront with other parties, especially if no business relationship is established after all. This information is not limited to the requested product or its specification but also includes other sensitive criteria (e.g., their price expectations). Likewise, buyers cannot tolerate any linking of their queries. Overall, buyers are concerned with leaking valuable data, fearing a loss of their competitiveness.

**G-I2: Seller Privacy.** Even though buyers are more likely to share information upfront (cf. **G-I1**), sellers also have an interest in confidentiality. Today, their privacy is at stake in two ways [AN07]. First, sellers reply to purchase inquiry requests with specific offers. Apart from the effort invested in this possibly unrewarding task (if no sale is closed), made offers reveal a variety of sensitive details. For example, they contain the company's production capabilities, its declared price, and potentially additional insights into available production resources or schedules. Thus, competitors might be able to derive the sellers' profit margins or other valuable details, eventually providing them with means to undercut offers. Second, currently, sellers might resort to publicly announcing their catalog (e.g., as known from consumer mail-ordering businesses) to attract business, i.e., their privacy is similarly affected.

**G-I3: Protocol Resistance.** Improved privacy during procurement further demands secure approaches, i.e., colluding parties should not be able to extract any additional information, especially about third parties. Similarly, the result of any privacy-preserving purchase inquiry must be sound, i.e., no manipulation must be possible at any time. This goal primarily concerns the comparison of price expectations, as otherwise, sellers could pretend to provide goods at every price to extract the requesting buyer's price limit. Furthermore, falsely-advertised products by a seller would be immediately noticeable to the buyer as the subsequent negotiation would fail right away (the "matched" seller cannot provide said products). Regardless, such unsound results would also diminish the value of and the trust in the protocol, i.e., sensible approaches should prevent such attacks to ensure technology acceptance.

**G-I4: Applicability.** Given our focus on real-world settings, any approach must satisfy the constraints of real-world applications. Namely, this goal covers both performance and scalability, where scalability refers to the number of products that are globally comparable and the number of contactable sellers in a specific period. Solutions not fulfilling these goals might only be able to improve the companies' privacy while failing to significantly improve the status quo.

Since purchases are started well in advance (usually providing several weeks of buffer), mainly due to the manual effort that is needed and to account for delivery delays, having an (automated) protocol run conclude within a single day constitutes a reasonable upper bound. If conducting a privacy-preserving purchase inquiry is not feasible, the envisioned dynamic business relationships remain impractical as they would incur significant overhead. Thus, to profit from all benefits, e.g., the ability to react to customer change requests, improved product quality, and lower costs, any proposed approach must scale to real-world needs. Thus, this goal is key for any solution's success and its technology acceptance (in industry).

**G-I5: Ease of Use.** To stress parts of the previous goal (**G-I4**), we explicitly model the ease of use as a distinct goal. In particular, we demand the independence of the involved parties, i.e., a purchase inquiry should not be bound to a fixed set of potential sellers. Instead, buyers should be able to contact sellers on demand and as needed, e.g., if no satisfying match has been made or the subsequent negotiations cannot be concluded. On a related note, to ease real-world deployment and since companies are reserved to use centralized services to manage their purchase inquiries [CGJ+09], newly-proposed designs should avoid them. All in all, this addition aligns nicely with **G-I3**, which mandates preventing all sorts of information leaks, because it effectively limits the number of involved stakeholders in a single protocol run to a minimum, i.e., by excluding uninvolved parties. Moreover, a direct bilateral protocol between a buyer and a single seller most likely reduces the load on the sellers (cf. **G-I4**) as they are only contacted if needed. Consequently, we argue that approaches to improve the privacy in purchase inquiries should avoid a trusted third party (to limit the threat of data leaks) and should not make use of multi-party computation (to improve the flexibility and to avoid round-based runs with fixed sets of potential sellers), i.e., we call for flexible, bilateral approaches.

**Non-Goals.** For this contribution, we focus on approaches that allow buyers and sellers to extend their established network of business relations in business-to-business markets without fearing any leaks of sensitive information. In particular, we do not want to replace procurement processes, contract negotiations, or approaches for bidding on products or prices in any way. Consequently, fuzzy queries are uncalled-for as buyers know the product properties that they inquire about. Instead, we intend to augment these established approaches with an intermediate step to establish new business relationships without the need to reveal sensitive information upfront.

#### 4.2.1.3 Related Work

While research efforts from different domains, including business experts and computer scientists, target the directions of electronic markets [KB06], auctions [NPS99,

NT00,Bra06], and private e-tendering [PPW08], they neglect to significantly improve the privacy during traditional procurement processes that are widely in use today. Consequently, companies protect themselves by limiting their considered sellers in practice [AN07] or requiring third parties to sign NDAs at an early stage of the negotiations [WR17], which is a time-consuming and tedious task (cf. Section 3.2.2).

Basically, prior work focuses on approaches that also cover the conclusion of purchases and sales, including agreeing on a fixed price. Thereby, they exceed our scenario, where only a pre-selection (matching) is needed to allow companies to still negotiate final prices. Furthermore, these works usually assume that one entity reveals what they are buying or selling, i.e., commonly, sellers have to reveal their product catalogs or products. Thus, these scenarios all contradict our goal of not sharing any information upfront as they only consider the privacy of one party.

In the area of advertising, related work [GCF11, HHHB11] only focuses on the privacy of ad receivers (cf. **G-I1**) as the second party, i.e., the ad provider, does not demand specific privacy guarantees. Thus, these approaches violate the need for seller privacy (**G-I2**). Regardless, they intuitively highlight similar privacy issues in other settings. To conclude, all of them fail to protect the privacy of all involved entities as they commonly focus on a single party only. Hence, they are not suitable.

Finally, initiatives such as the federated data infrastructure GAIA-X [BFRLG21] and IDS [OAC+16, OJ19] aim at standardizing (industrial) data sharing. However, they are broader in scope, impose significant technical and organizational requirements for participating, and so far have not specifically addressed purchase inquiries. Simply applying other approaches for the matching and exchange of information to our considered scenario is not feasible either because, for purchase inquiries, multiple dimensions must be compared jointly (i.e., product(s) and price range(s)). Consequently, we stress the research gap of offering privacy-preserving purchase inquiries through technical and secure approaches that scale to industry needs.

### 4.2.2 Two-Way Privacy for Purchase Inquiries

For our second contribution, we select several building blocks from privacy-preserving computation (cf. Section 2.3.1), i.e., PSI and HE, to propose multiple designs that ensure two-way privacy for purchase inquiries. These designs differ in terms of their confidentiality guarantees and the computational overhead they introduce for the involved buyers and sellers. Hence, companies can choose a design according to their use case-specific needs. Overall, we provide the evolving industrial landscape with an important tool that promises the risk-free establishment of dynamic buyer-seller relationships. In principle, buyers can reliably increase the number of considered sellers as they do not have to fear any disadvantages from sharing sensitive information upfront (with unsuitable or untrusted sellers). Our proposed designs have in common that they are neatly integrateable into today's well-established procurement processes. Thus, the other stages of procurement processes remain unaltered, eliminating any repercussions that follow from the use of privacy-preserving purchase inquiries. Moreover, since they are oblivious of the handled products, their deployment is independent of concrete domains, allowing for industry-wide use.

In the following, in Section 4.2.2.1, we give an overview of our designs before detailing their respective protocols in Section 4.2.2.2. Afterward, in Section 4.2.2.3, we introduce our prototypical implementations. Since their feasibility for real-world deployments is of utmost importance, we then proceed with an extensive evaluation that covers four subsubsections: In Section 4.2.2.4, we first discuss the designs' general performance. Subsequently, in Section 4.2.2.5, we specifically look at real-world use cases to underline our work's feasibility. In addition to the performance, we need to carefully assess the security and privacy guarantees of our designs to ensure that our work improves the situation for companies in the IIoT. Thus, in Section 4.2.2.6, we elaborate on these matters. Finally, we conclude our evaluation in Section 4.2.2.7.

#### 4.2.2.1 Designs for Secure and Privacy-Preserving Bilateral Purchase Inquiries

We now propose our designs. As a foundation, we first introduce a semi-formal definition of purchase inquiries along with the properties of each involved party. Sourcing this notation, we then continue with a design overview.

##### Notation for Purchase Inquiries

While our definition focuses on a single Buyer $B$ who considers $n$ potential Sellers $S_1, \ldots, S_n$, our proposed, bilateral protocols only handle a single seller $S_i$ at a time. We can handle multiple independent buyers by parallelizing protocol runs conveniently. Thus, corresponding designs scale to any real-world setting at hand.

**Product Modeling.** Every relevant product $P$ is representable by a unique identifier. To this end, for discretization, we define a global (domain-specific) modeling function $f : \mathbb{P} \longrightarrow \mathbb{N}, \forall P \in \mathbb{P} : f(P) \in [0, \ldots N]$, where $\mathbb{P}$ matches all relevant products, $N \geq |\mathbb{P}|$ is a fixed integer and globally defined together with $f$ for a specific domain. In a query, buyers and sellers must use the same $f$ to ensure interoperability.

**Product-Price Mapping(s).** First, we define a set $X = f(P) \times \ddot{\mathbb{M}}$, where $\ddot{\mathbb{M}}$ refers to a price, e.g., in USD. A tuple in $X$ semantically corresponds to $(id, price)$. We define a Buyer $B$'s query $q$ as $P_q^B \subset X$, and for every Seller $S_i$, we define a set containing all producible items in her product catalog $c$ as $P_c^{S_i} \subset X$. As we require specific price expectations per product, we further note that $\nexists (id, m_1), (id, m_2) \in P_q^B \wedge \nexists (id, m_1), (id, m_2) \in P_c^{S_i}$ where $m_1 \neq m_2$. Buyers and sellers independently populate their sets $P_q^B$ and $P_c^{S_i}$: For each product $P$ with $id$, $max_{id}^B$ defines the maximum *price* a Buyer $B$ is willing to pay for it, and $min_{id}^{S_i}$ indicates the minimum *price* expected by Seller $S_i$. For eventual sales negotiations, the respective *min* and *max* values are kept private.

**Buyer.** A Buyer $B$'s query $q$ is expressed through the set $P_q^B$, where to-be-queried products are stored in direct connection with their envisioned maximum prices. We further define a function $g_B$ that indicates whether Buyer $B$ is interested ($1 \equiv true$) in a specific product, i.e., $g_B : \mathbb{N} \longrightarrow \{0, 1\}, \forall id \in f(\mathbb{P}) : g_B(id) = 1 \Leftrightarrow (id, m) \in P_q^B$, else $g_B(id) = 0$. For each query $q$, we further compute a specific price threshold $\perp_B = \sum_{id \in f(\mathbb{P})} g_B(id) \cdot max_{id}^B$ to fix the maximum costs.

**Seller.** In addition to the product catalog $c$ and the conceived minimum prices that are expressed through $P_c^{S_i}$, we define a globally-defined price threshold $\top \notin \ddot{\mathbb{M}}$. We rely on $\top$ as a price placeholder for every product $id$ that is not listed in the seller's product catalog $c$, i.e., $(id, \top) \in X \Leftrightarrow \nexists (id, m) \in P_c^{S_i}$.

**Purchase Inquiry.** We express a purchase inquiry $PI$ between Buyer $B$ and Seller $S_i$ as $PI_i(P_q^B, P_c^{S_i})$. Further, $PI$'s result is either: (i) In a more expressive (fine–granular), yet more revealing $PI$, Buyer $B$ obtains a result for each queried product in her query $q$: $\forall\,(id, price) \in P_q^B : (id, price, \{0, 1\})$, or (ii) she only learns a single result 1 or 0 (indicating a match or no match, respectively), stripping all details.

When summarizing this notation informally, we identify two key aspects. First, Buyer $B$ has a query $q$ that contains different, parameterized products as well as a maximum price for each of these products. Second, each Seller $S_i$ maintains a product catalog $c$ with producible items and corresponding minimum prices.

### Design Overview

We name our designs according to their underlying concepts. First, we are proposing a PSI-based design, called PPI, that utilizes two computational phases to process a purchase inquiry. Second, our primary HE-based design, called HPI, builds on computations on encrypted inputs. Third, based on HPI, we derive a design variant that obliviously offloads parts of expensive HE computations to a cloud. Accordingly, we refer to this cloud-tailored design as cHPI. Mixing or combining our designs as part of a single purchase inquiry between buyer and seller is not possible. However, a Buyer $B$ can select different designs for different Sellers $S_i$ because our bilateral realization ensures that individual runs are independent of each other.

We summarize the differences between our designs in more detail in Section 4.2.2.7, but first, now in brief: Due to its simplicity and low-overhead building blocks, PPI promises improved performance over HPI and cHPI at the expense of slightly weaker privacy guarantees. In contrast, HPI and cHPI primarily differ in where they distribute the computational overhead of performing computations on HE ciphertexts.

Our designs to realize *privacy-preserving* purchase inquiries consist of four phases, as we illustrate in Figure 4.10. Initially, ① as part of a global setup, a global modeling function $f(P)$ must be defined, which is later used to deterministically map products to $id$s. Next, each buyer-seller pair ($B$, $S_i$) bilaterally executes the protocol steps, i.e., to consider multiple potential sellers, the buyer reruns the protocol multiple times. In the first protocol step, ② buyer and seller can individually pre-process certain steps, such as preparing the query $q$ or indexing the catalog $c$. The exact opportunity for this pre-processing depends on the specific design in use.

Afterward, ③ we *bilaterally* conduct the privacy-preserving comparison phase to obliviously identify matches (in terms of product(s) and price range(s)) between the buyer's query and the seller's catalog. For our straightforward, PSI-based purchase inquiry design (PPI), we utilize a two-phased approach: an independent matching and a subsequent price comparison for each matched product. In the second phase, we have to outsource the ORE-based price comparison to a third party to preserve

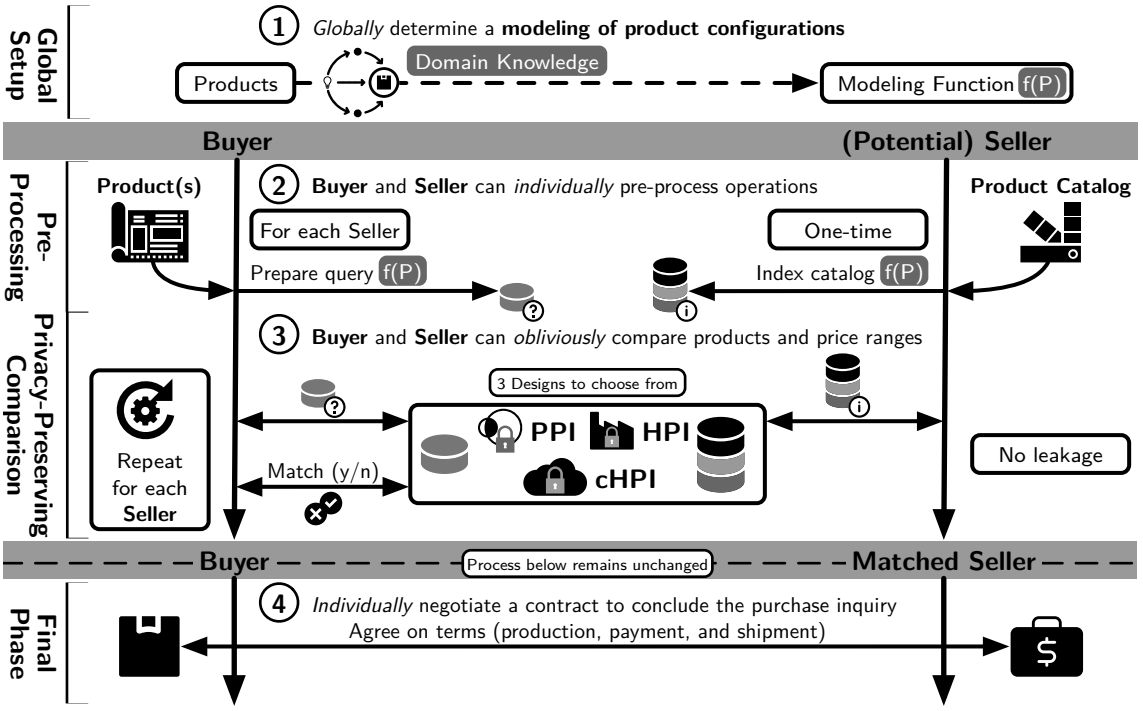



**Figure 4.10** Conceptually, our designs consist of phases to realize a privacy-preserving comparison of a buyer's query and a seller's catalog: They depend on a global modeling, allow for pre-processing, and compute the comparison. Finally, after a successful match for a query (in terms of product(s) and price range(s)), the traditional procurement process can continue.

privacy as ORE supports symmetric-key cryptography only. Thus, PPI partially violates the ease of use goal (**G-I5**). In contrast, our HE-based approaches (HPI and cHPI) directly return a single result for all queried products following the computation. The buyer repeats this phase for each potential seller. Eventually, she knows which sellers are, in principle, able to satisfy her requested order within the expected price range. These two aspects are the most important decision factors [XCK08].

Due to its more performant PSI-based building block, PPI outperforms our HE-based designs HPI and cHPI (cf. Section 4.2.2.4). However, its superior flexibility and performance come at the expense of slightly-weaker privacy guarantees due to its straightforward design, as we detail in Section 4.2.2.6. To relieve the seller in HPI from some computational load, we can offload parts of the (costly) computation to an untrusted cloud, which is inspired by work that outsources vector multiplications [CG17]. Depending on the seller's computing and networking resources, cHPI can be a suitable alternative as it reduces her computational load at the expense of greatly-increased network traffic. While such a design seems to contradict **G-I5**, using HE ensures that the cloud learns nothing about the data it operates on nor the final result. Consequently, cHPI does not require any trust in this third party.

Finally, ④ the buyer can contact any number of matched sellers to continue with the final negotiations, i.e., to agree on a price, a delivery schedule, and other relevant aspects. This final phase, after our privacy-preserving comparison, remains unchanged regarding established procurement processes. Thereby, both buyers and sellers keep the same flexibility as they are accustomed to today.

Our designs generally ensure buyer and seller privacy (**G-I1** and **G-I2**) by utilizing building blocks from privacy-preserving computation (cf. Section 2.3.1). Moreover, thanks to our modular phases, we can gradually adjust and tune specific parts of each protocol if needed. As a protocol run only concerns the buyer and a specific seller, we also account for protocol resistance (**G-I3**), unlinkability of buyer queries (**G-I1**), and the desired ease of use (**G-I5**). Furthermore, with our non-invasive impact on today's procurement, our privacy-preserving purchase inquiries integrate neatly into established businesses processes: We allow buyers to consider a larger set of potential sellers and relieve sellers from the need to draft offers early on while removing the need to disclose any sensitive information upfront for all parties. The exact contract negotiations (with optional soft criteria) are part of subsequent procurement steps.

#### 4.2.2.2 Specifics of the Protocol Phases

Now, we discuss the phases of our proposed designs in more detail. We follow the logical flow from Figure 4.10 and include protocol-specific sequence charts.

##### Modeling of Product Configurations

In the first phase, we obtain a unique product identifier from a parameterized product description. Subsequent phases only make use of these discrete product identifiers. Rather than an algorithm that computes them from a given description, we present a data governance process that leads to machine-comprehensible product descriptions with unique identifiers—an approach widely used, e.g., in e-commerce in context with the schema.org product description schema (cf. Section 4.2.2.3). This global and deterministic approach also conveniently enables additions, refinements, updates, and product deprecations. The input into this process captures all relevant product information except for the price. The output is a graph-shaped representation in a uniform terminology, where the graph's root node carries the product's identifier.

To ensure global access to the market for all procuring companies, product identifiers, descriptions, and purchase inquiries need to be FAIR [WDA+16], i.e., findable, accessible, interoperable, and reusable. While "I" and "R" directly apply in our scenario, "F" and "A" are only desirable in some use cases, e.g., a seller maintaining a public catalog. Technically, FAIR data is often implemented along with the principles for 5-star open data [5st12, HRS18] and linked data [BL06, HAPS13], with an open license being an optional concept: uniform resource identifiers (URIs) are used as globally unique identifiers of things (products, standards, etc.), and data is linked to other data to provide context, e.g., "what standards does product $P$ conform with", or "how is the 'weight' of a product defined". 5-star data implementations usually adopt the graph-based RDF [WLC14] data model, which natively uses URIs for instances, e.g., products, as well as on the schema level, e.g., to define a property *weight*, as we exemplarily illustrate for a product in Figure 4.11.

Schema terms are usually agreed upon by a larger community in the domain (often moderated by a standardization body). Such a formalized schema is called *ontology*. Thus, a product $P$ is modeled as an RDF resource, i.e., a node in the graph that has

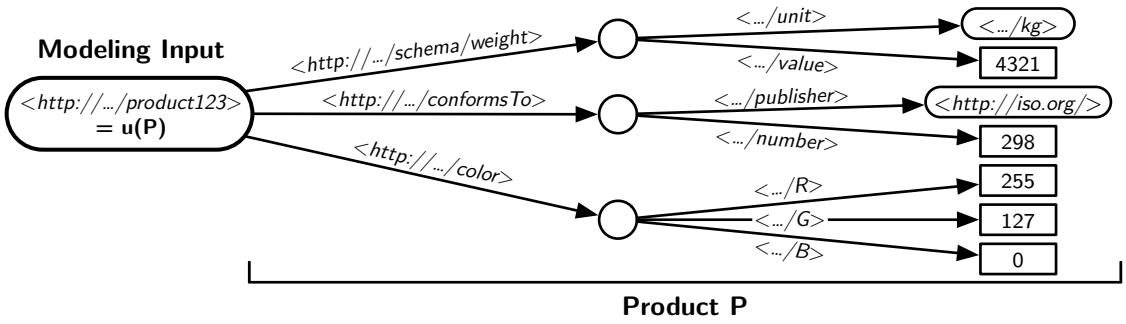



**Figure 4.11** With our modeling of product $P$, we are able to obtain a deterministic representation. Here, we exemplify such a product description, which is illustrated an RDF graph.

a description in terms of outgoing edges pointing to further nodes. Given a URI $u(P)$ of $P$, which is unique and discrete, $f(u(P))$ maps it to a positive integer, as required by our designs. Henceforth, we abbreviate it as $f(P)$. As we lack a universal, canonical definition of how to generate URIs for resources with RDF, future work has to also agree on globally-defined data governance (or implementations).

### Pre-Processing Opportunities

After the global definition of the modeling function $f(P)$ to discretize products $P$, we now present our purchase inquiry protocols. Given that the pre-processing opportunities vary between our designs, we now discuss them individually.

**PPI: A Two-Phased Approach.** We illustrate this design in Figure 4.12. ① For the initial comparison of PPI (the matching step), Seller $S_i$ can pre-generate the Bloom filter (containing all offered products $P_c^{S_i}$) that is used for the PSI. Using $f(P)$ and RSA blinds [KLS+17], the potential seller inserts all computed product *id*s in the Bloom filter. Buyers later receiving the Bloom filter cannot brute-force the inserted products, as all inserted product *id*s are signed using the seller's private key. While this step is the most computationally-expensive task, it still is reasonable as the seller only has to perform it once. PSIs that build on other technologies [MAL23], e.g., HE or OTs, do not necessarily provide comparable pre-processing capabilities for our design. If desired, Seller $S_i$ can create a new Bloom filter by re-generating the RSA key to protect against buyers colluding to derive her producible items, i.e., the respective *id*s, from $P_c^{S_i}$ that match to her product catalog $c$. Otherwise, multiple buyers could later merge their queries, as the seller RSA-signs the queried *id*s according to the RSA-PSI protocol [KLS+17] using the same private key. This pre-processing, including the RSA blinding, exploits the unmodified RSA-PSI protocol by Kiss et al. [KLS+17] for efficiency. We simply outsource this step from the originally-proposed protocol sequence.

To prepare a specific query, Buyer $B$ can also pre-process the (quick and inexpensive) derivation of product *id*s in $P_q^B$ using $f(P)$ for the subsequent comparison.

**HPI: An HE-Based Protocol.** In Figure 4.13, we present this design's sequence chart. ① Buyer $B$ homomorphically encrypts the result of $g_B(id)$ for every product

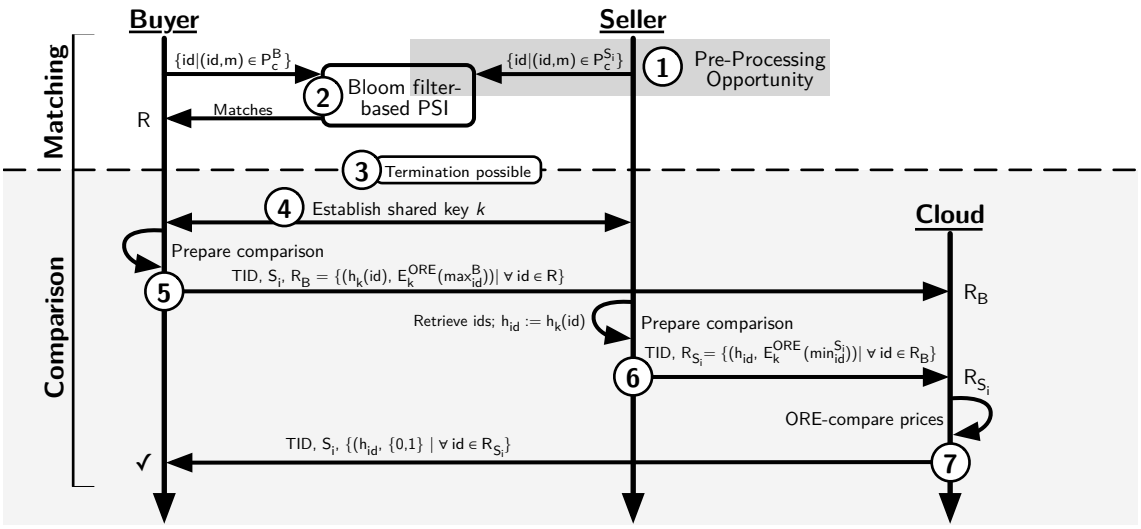



**Figure 4.12** PPI consists of two consecutive computational phases. After identifying matches of product *id*s with the contacted seller during the PSI-based matching phase, the buyer can trigger the cloud (third-party) to obliviously conduct an ORE-based price comparison.

$id$ to pre-process a query, creating $C_B$ for subsequent use in HPI. Additionally, she can pre-compute the query-specific price threshold $\perp_B$ and PHE-encrypt it. To prevent correlation attacks among contacted sellers, she can prepare individual PHE ciphertexts for each Seller $S_i$. Alternatively, she can reuse them. (2) Each potential seller $S_i$ can already blind her prices from $P_c^{S_i}$ to compute $M_{S_i}$ for subsequent use in HPI. However, this step is not computationally expensive as no encryption or signing is required. After this pre-processing phase, using the modeling function $f(P)$, the sets $P_q^B$ and $P_c^{S_i}$ have been individually prepared for the purchase inquiry $PI_i(P_q^B, P_c^{S_i})$ by Buyer $B$ and Seller $S_i$.

**cHPI: A Cloud-Tailored Design Variant.** As we detail in cHPI's sequence chart in Figure 4.14, the (1) pre-processing of cHPI is identical to the pre-processing of HPI.

### Privacy-Preserving Comparison

Subsequently to the parties computing their query and product catalogs, the buyer can initiate the inquiry with each seller, i.e., she can trigger respective protocol runs for as many sellers as needed, e.g., until a suitable seller has been found.

**PPI: A Two-Phased Approach.** In Figure 4.12, we illustrate the two individual steps of PPI, i.e., matching and comparison. (2) The matching step continues the Bloom filter-based PSI protocol by Kiss et al. [KLS+17]. First, Buyer $B$ blinds her $P_q^B$ using the public RSA key of Seller $S_i$ and sends the blinded query to the seller. Then, the seller signs these entries using her private RSA key before returning the results to the buyer. Lastly, in the matching step, the buyer can remove the respective blinds and check for containment in the Bloom filter, which concludes the matching. For each product, the buyer learns whether the seller is able to produce it or not. Now, (3) the buyer can gracefully terminate PPI, e.g., if the result indicates no or only partial matches of the requested products.

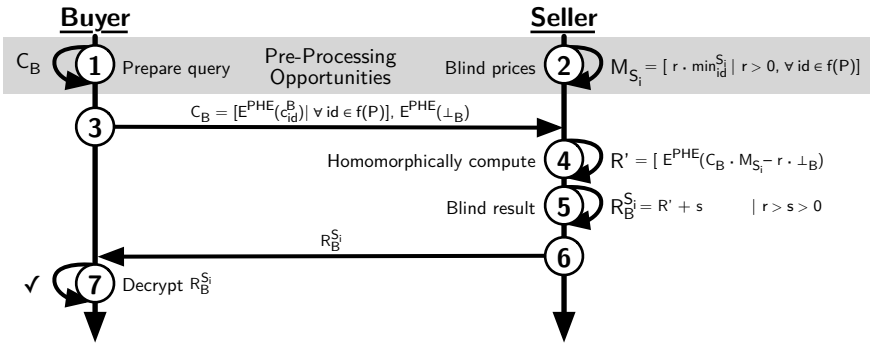



**Figure 4.13** After the pre-processing in HPI, the seller homomorphically computes a blinded result and returns it to the buyer for decryption. This result contains the inquiry's outcome.

When continuing with the purchase inquiry, Buyer $B$ can subsequently trigger the comparison step of PPI. Initially, ④ Buyer $B$ and Seller $S_i$ establish a shared key $k$ that is used to order-revealingly encrypt their price expectations. Using a cryptographic, keyed hash function $h_k$, they further obfuscate the matched product $id$s. In line with our ease of use goal (**G-I5**), we consider involving a third party as disadvantageous. Still, in PPI, using a random transaction identifier $TID$, the buyer must task an independent third party in the cloud to compare the prices under ORE. To this end, the cloud explicitly queries the (encrypted) prices from Seller $S_i$. First, ⑤ the buyer transmits $R_B$, a set containing the obfuscated $id$s and her ORE-encrypted price expectations, to the cloud, which requests the seller's price expectations. Second, ⑥ the seller shares $R_{S_i}$, which likewise contains the obfuscated $id$s and her ORE-encrypted price expectations, with the cloud to allow for an ORE-based price comparison. Finally, ⑦ Buyer $B$ receives a result for each queried/matched product from the cloud, revealing whether the price expectations overlap as well.

**HPI: An HE-Based Protocol.** We detail the steps of HPI in Figure 4.13. First, ③ Buyer $B$ transmits a PHE-encrypted vector $C_B$, containing encrypted zeros and ones from $g_B(id)$ and her encrypted price threshold ($\perp_B$), to Seller $S_i$. Subsequently, ④ Seller $S_i$ can homomorphically compute the result of the purchase inquiry using her blinded prices ($M_{S_i}$), i.e., she multiplies $C_B$ with her product prices (as a scalar product). Afterward, she blinds $\perp_B$ and subtracts it from the scalar product to obtain $R$'. ⑤ She blinds the intermediate result ($R$') before ⑥ returning the blinded result $R_B^{S_i}$ to the buyer. Finally, ⑦ the buyer can decrypt the PHE ciphertext ($R_B^{S_i}$).

If the result is smaller than 0, the seller can produce the queried items and offers them in the desired price range. Otherwise, the contacted seller is not a match. In contrast to the default behavior of PPI, in HPI, the buyer gains no knowledge about individual product matches in her query $P_q^B$, improving the seller's privacy (**G-I2**).

**cHPI: A Cloud-Tailored Design Variant.** cHPI largely follows the same computational concept as HPI. The buyer sends her encrypted vector directly to the cloud, and the seller shares her blinded prices with the cloud. For confidentiality, these inputs are blinded with a random vector that buyer and seller agreed upon. The cloud then computes the (costly) scalar product before returning the result to the seller (to allow for blinding, as in HPI). We proceed with a detailed discussion of cHPI and illustrate the computation steps of this cloud-tailored variant in Figure 4.14.

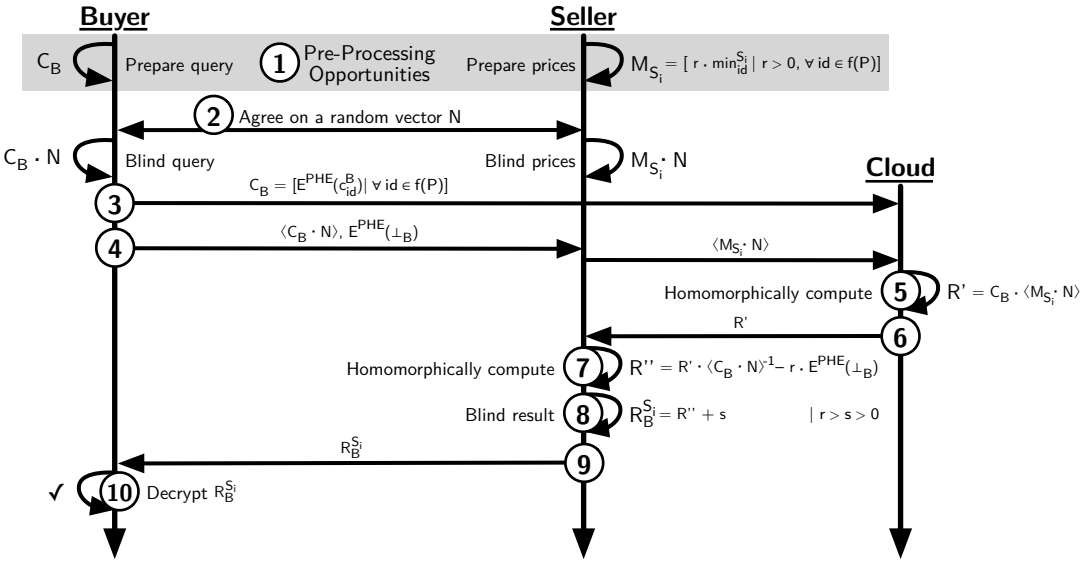



**Figure 4.14** In comparison to HPI, the seller offloads the computationally-expensive homomorphic computation to a cloud in cHPI, which significantly reduces her computational workload.

Following the pre-processing, ② Buyer $B$ and Seller $S_i$ agree on a random vector $N$ to blind their pre-processed values ($C_B$ and $M_{S_i}$). Otherwise, the cloud would know the seller's prices as they are not encrypted in HPI due to the local computation (no threat of information leakage). Afterward, Buyer $B$ must share her encrypted inputs with both ③ the cloud ($C_B$, unblinded vector) to eventually allow for a removal of the blinds from the cloud-computed result by the seller and ④ the seller ($\langle C_B \cdot N \rangle$, blinded scalar product). ⑤ Using the seller's blinded prices ($\langle M_{S_i} \cdot N \rangle$), the cloud can homomorphically compute the (computationally-expensive) scalar product, i.e., $R$' (as in HPI). Then, ⑥ the cloud returns this result ($R$') to the seller, who ⑦ removes the random vector $N$ using the blinded scalar product $\langle C_B \cdot N \rangle$. Finally, as the only local computation step, the seller subtracts the blinded $\perp_B$ to compute the intermediate result $R$". ⑧ She further blinds the result with $s$ to obtain $R_B^{S_i}$, and ⑨ returns this single result to the buyer, who ⑩ can decrypt the result using her private PHE key. In cHPI, the result's semantics are identical to the ones in HPI.

After this phase, the purchase inquiry $PI_i$ is concluded, and the buyer is aware of the result(s), i.e., whether her query (including the price thresholds) fits the sellers' catalog and price expectations. The privacy-preserving comparison is oblivious in our designs, i.e., no sensitive information is leaked or exchanged. In contrast to PPI, where the buyer is aware of the matches for each product, in HPI and cHPI, the buyer only learns whether a seller can produce all requested products within the specified price range. We discuss the corresponding implications in Section 4.2.2.6.

### Concluding the Purchase Inquiry

In the final phase, the buyer can contact one or multiple matched sellers to individually continue with the procurement process. In particular, they have to negotiate contracts, i.e., agree on specific terms. The buyer is also able to include additional aspects, such as sustainability or reputation, into the decision-making when contacting

matched sellers. The introduction of our privacy-preserving purchase inquiry does not affect other steps of the procurement process. Thus, we consider the next steps as out of scope and refer to related work (e.g., [BTHN96, BVC01]). The conducted (privacy-preserving) comparison is an indicator that, in theory, an agreement can be reached. By design, buyers and sellers can (still) freely negotiate prices and terms.

#### 4.2.2.3 Real-World Realization

In the following, we present our Python-based prototypes. Our designs build on well-known building blocks by linking their (secure) operations to preserve the participants' privacy. Overall, the seller and cloud components provide RESTful APIs through Flask [Ron10] web servers that offer secure network connections. Furthermore, to manage all tasks, e.g., incoming queries, we utilize Celery [Sol09]. To account for numerous requests in parallel, we support a separation of frontend (API) and backend (workers). Thus, companies can scale their resources as needed, e.g., by relying on cloud computing to offload computationally-expensive tasks.

**Modeling of the Product Configuration.** For our evaluation, we rely on a simple discretization approach as modeling (cf. Section 4.2.2.5). For real-world use, we would have to define a way of generating ("minting") URIs to identify the things described [SC08] (cf. Section 4.2.2.2). In this regard, market participants are free to agree on either descriptive URIs (`https://trusted-marketplace.com/machine-tools/dmu-50-gen3`) or non-descriptive URIs (`https://vdma.org/id/23597656-0e29-4a04-9f6f-0a8d856c3769`), as long as it has been agreed upon how to retrieve information about a thing, given its URI. For example, following pure linked data best practices does not require directory services but requires URIs to be HTTP URLs, from which RDF metadata is downloadable. While we consider the exact realization as orthogonal research that is independent of our designs, a trusted party (e.g., an industry association), could maintain these globally used URIs on behalf of the products' providers within an independent Internet domain to also ensure the privacy of retrieving companies. The RDF data should use agreed-upon schemas. Thus, we argue to prefer re-using existing schemas, e.g., ECLASS for product classes and product properties [ECL07], for which an old, unofficial ontology exists [HR10], and an official one is under development, or many advanced units of measurement ontologies to choose from [KS19]. Despite being less widespread, unofficial ontologies also exist for describing standards [BGGN+20]. When ontologies for use with $f$ are missing, we recommended adapting existing ones, e.g., specializing the general, domain-independent GoodRelations ontology for product descriptions in e-commerce (now part of the search engine standard schema.org [Hep15]), or generalizing ontologies for specific machine tool applications, e.g., ExtruOnt [RDBI20].

**PPI: A Two-Phased Approach.** We rely on a Bloom filter-based PSI [KLS+17] for the matching step as this variant promises to be very performant (resulting in fewer round trips and the ability to pre-process the product catalog). We utilize a Python library [Ben20], which is based on PyCryptodome [Eij14], and added (de)serialization support. For the comparison, we use a Python library [Pat17], which implements the ORE scheme by Chenette et al. [CLWW16].

**HPI and cHPI: Our HE-Based Protocols.** The remaining designs build on PHE (cf. Section 2.3.1). We rely on the Paillier cryptosystem [Pai99]. More specifically, we use CSIRO Data61's Python library [CSI14] for our implementation.

Our implementations of PPI, HPI, and cHPI are publicly available [SrcC23a].

#### 4.2.2.4 Performance Evaluation

As we have detailed in Section 4.2.2.1, our designs fulfill the conceptual goals of buyer privacy (**G-I1**), seller privacy (**G-I2**), and ease of use (**G-I5**). In the following, we now take a look at the performance and scalability of our designs (**G-I4**). In particular, we present our experimental setup and discuss our measurements of the respective runtimes, memory usage, and network load. Finally, based on these results, we compare the different designs with each other. After this synthetic evaluation, we study our designs in light of two real-world applications in Section 4.2.2.5.

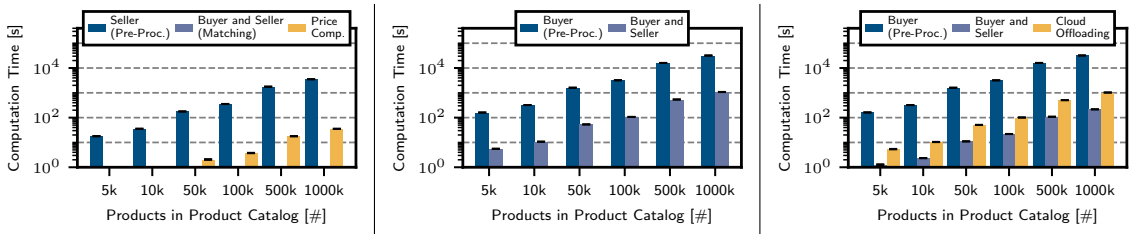
##### Experimental Setup

We utilized a single server (Intel Xeon E5-2630 with 32 GB RAM) to host all involved parties. We ran Docker containers [Mer14], and they communicated via the loopback interface. We report on the arithmetic mean of 30 runs, calculate 99 % confidence intervals, and present all measurements in logarithmic scales. When evaluating PPI, we configured an RSA key size of 2048 bit for the PSI, and the defined key size of the hash function in ORE was 20 bit. For the Paillier cryptosystem, in HPI and cHPI, we relied on a key size of 3072 bit to encrypt the PHE ciphertexts.

##### Performance Measurements

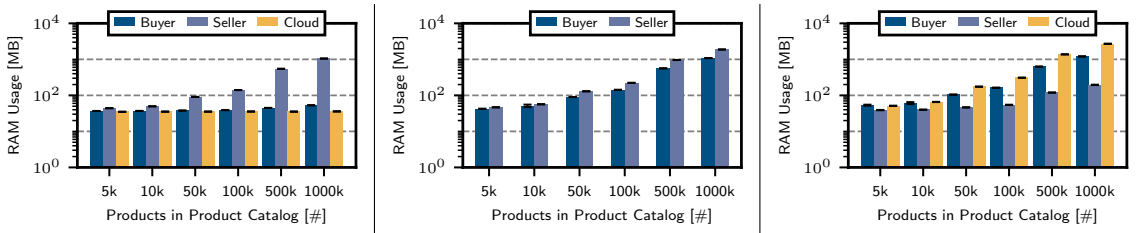
We measured the computation time, maximum RAM usage, and the size of outgoing transmissions per party. As the number of products in the product catalog significantly influences the performance, we consider different, real-world-derived values between 5000 and 1 Mio. for our evaluation. Contrary, the query size only has negligible influence on the runtime during the PSI-based matching as the buyer has to request an encryption for each product (in PPI) or no influence at all as every product *id* has to be encrypted anyway (HPI and cHPI). Given that $P_q^B \ll P_c^{S_i}$ and real-world queries cover only a few products, we fix the query size at 10.

**PPI: A Two-Phased Approach.** As we present in Figure 4.15a, the runtime of PPI increases linearly with the product catalog's size due to the signed values used in the Bloom filter-based PSI (cf. pre-processing opportunities in Section 4.2.2.2). Notably, it is shaped by the seller's pre-processing, i.e., the runtime is dominated by a phase that is only needed before the first protocol run and can be omitted by subsequent runs with different buyers. The remaining runtime splits between the PSI during the matching step (negligible: 500 ms with a product catalog of 1 Mio. entries and 10 products in the query) and the ORE-based price comparison computed in the cloud, which sequentially involves buyer and seller. In the latter step, the most workload




**(a)** Runtimes of PPI **(b)** Runtimes of HPI **(c)** Runtimes of cHPI

**Figure 4.15** The runtime of our designs scales linearly with the number of products in the product catalog. PPI is faster than HPI and cHPI by one order of magnitude. While PPI allows for significant seller pre-processing, HPI enables buyer pre-processing. cHPI reduces the seller's workload by one order of magnitude in comparison to the purely local HPI protocol.




**(a)** Memory needs of PPI **(b)** Memory needs of HPI **(c)** Memory needs of cHPI

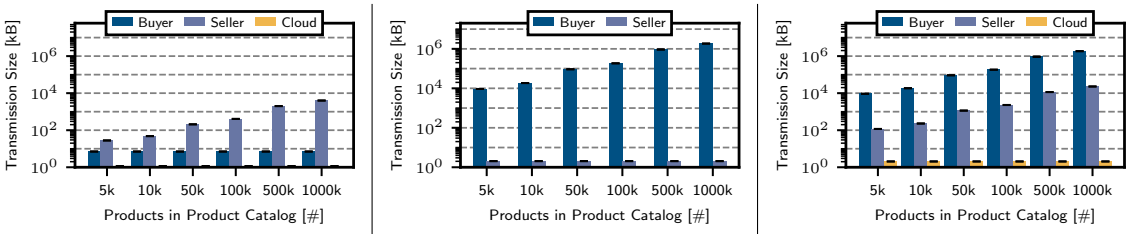
**Figure 4.16** In PPI, the seller keeps the entire Bloom filter in memory, which conforms to the linear increase of memory usage. HPI reveals comparable memory needs for buyers and sellers, while the cloud in cHPI significantly reduces the seller's memory usage when compared to HPI.

is on the seller (36 s with 1 Mio. total products) as she has to hash all products with the shared key $k$ to identify the cloud-requested prices. In contrast, as part of the price comparison, the cloud is only affected by the number of overall matches. For example, with 10 matches, we measure 75 ms of workload in the cloud.

The memory needs of PPI (Figure 4.16a) are low. While the cloud's use is independent of the catalog's size (at most 35 MB), the memory consumption at the buyer correlates with the size of the Bloom filter (4 MB with 1 Mio. products). The seller's memory usage during the matching is only 0.5 MB and increases linearly with the catalog's size during the price comparison (262 MB with 1 Mio. products). Most memory is needed for the pre-processing (at most 1.9 GB for 1 Mio. products).

Similar observations also hold for the transmission sizes (Figure 4.17a). For a query with 10 products (and in our setting 10 matches), buyer and cloud transmit 7.1 kB and 1.2 kB, respectively. The size of the transmitted data by the seller increases linearly but remains maintainable (only 4 MB with 1 Mio. products in the product catalog). In general, with a total runtime of 59 min (including the one-time pre-processing), moderate RAM usage, and limited network transmissions even for product catalogs with 1 Mio. products, PPI is very suitable for real-world use. Sellers using pre-processed catalogs or buyers aborting the protocol whenever the matching step was unsatisfactory further reduce the overall resource needs of PPI.

**HPI: An HE-Based Protocol.** The runtime of HPI (Figure 4.15b) increases linearly with the number of products. While it is dominated by the buyer's pre-processing




**(a)** Network load of PPI **(b)** Network load of HPI **(c)** Network load of cHPI

**Figure 4.17** In PPI, the seller transfers the Bloom filter for the PSI-based matching, and buyer and cloud correlate with the query size and the number of matches, respectively. In HPI, the buyer has to upload the encrypted catalog and only receives a single result from the seller. In contrast to HPI, the seller offloads the PHE-based computation to the cloud when using cHPI.

(8.8 h for 1 Mio. products), the remaining privacy-preserving comparison, involving both parties, is nearly 30 times faster (≈ 18 min for 1 Mio. products). Hence, HPI primarily burdens the buyer. However, she can reuse every pre-processed query by sending it to all considered sellers and thereby significantly reducing her load.

In HPI, the maximal memory usage (Figure 4.16b) correlates linearly with the number of products as the buyer and seller load all ciphertexts into memory for the pre-processing and computation, respectively. The maximum lies at maintainable 1.1 GB for 1 Mio. products on the buyer's and 1.9 GB on the seller's side. Still, to reduce the memory usage resulting in a constant maximum, we could easily adjust our implementation to process the products (computation) in smaller batches.

Concerning the network transmissions (Figure 4.17b), the buyer has to upload all PHE ciphertexts. Thus, the amount scales linearly with the number of products (1.85 GB for 1 Mio. products). In contrast, the seller only returns a single result, i.e., we constantly measure only 2.1 kB. While the necessary resources are steep in HPI, its needs are still maintainable for real-world deployments. Today's non-automated procurement usually takes several days up to weeks. Thus, introducing a processing of 10 h for catalogs with 1 Mio. products adds no significant overhead. A RAM consumption of at most 1.9 GB and network transmissions of 1.85 GB do not overload today's infrastructures and the resources of companies either.

**cHPI: A Cloud-Tailored Design Variant.** In Figure 4.15c, we detail the overall runtime of cHPI. We notice that the offloading takes most of the runtime (excluding the buyer's initial pre-processing) in cHPI. Notably, in comparison to HPI (Figure 4.15b), the seller's workload is reduced by nearly one order of magnitude due to the offloaded scalar product computation relieving many of her resources. Thus, the offloading to the cloud significantly unburdens the seller as most workload is now shifted to the cloud. In contrast, the buyer's pre-processing remains identical as she still has to encrypt the same number of PHE ciphertexts when using cHPI.

Concerning the RAM usage, the seller is also relieved in cHPI as she no longer has to keep all ciphertexts in memory, as we detail in Figure 4.16c. Instead, the cloud inherits this task, effectively reducing the seller's needs when comparing HPI to cHPI. As for HPI, we could also adjust our cHPI implementation to support batch processing. Thus, we could end up with an upper bound for the cloud if needed.

Finally, as we have expressed before (cf. design overview in Section 4.2.2.1), we measure significant differences in network usage when comparing HPI to cHPI. As we illustrate in Figure 4.17c, the seller has the burden of offloading her blinded prices to the cloud when using cHPI, resulting in a transmission overhead when compared to HPI. Thus, in cHPI, the seller's network usage correlates with the number of products. This observation also holds for the buyer in both HPI and cHPI. In contrast, the cloud's network usage is constant as the result is always a single PHE ciphertext, which is irrespective of the number of products that were used as initial input. This behavior matches the seller's network usage in HPI. Depending on the exact scenario, the significantly-reduced computational and memory needs at the seller will outweigh the increased transmission overhead. Given that the overall runtime is still suitable for real-world use, even with constrained network links, e.g., a product catalog with 1 Mio. products incurs a seller upload of only 23 MB, cHPI is a promising design variant of HPI for real-world deployments.

#### Comparing the Performance of PPI to HPI and cHPI

The resource usage and distribution over all participants is crucial for real-world use of our proposed designs. Hence, in the following, we look into their respective differences. The main reason for PPI outperforming HPI and cHPI is the difference in the computation. While the HE-based designs, HPI and cHPI, operate with ciphertext of every product *id*, even if they are irrelevant to both buyer and seller. In contrast, PPI only operates with product *id*s that are relevant for the current purchase inquiry, i.e., the *id*s in $P_q^B$ and $P_c^{S_i}$. This conceptual difference follows from the operation of the underlying building blocks (and their confidentiality guarantees).

*Computation Time.* Given that the total runtime scales with the product catalog's size, all parties can estimate their computation time properly before starting the execution. While all designs are well feasible for all involved entities with a reasonable product catalog size, we notice that PPI is one order of magnitude faster than HPI and cHPI. This difference in runtime bases on the PHE-induced overhead in HPI and cHPI, which enables very strong privacy in comparison to the comparably-inexpensive PSI-based, Bloom filter-enabled matching in PPI. While the overall runtime of cHPI is identical to HPI in our measurements, in practice, the presented results would likely differ as the offloading in cHPI significantly increases the transmission size of the seller when compared to HPI, with most companies only having access to constraint network links. Irrespective of the overall runtime, all protocols enable specific entities to pre-compute parts of the protocol and reuse these parts over several runs. The quicker PPI protocol allows the seller to pre-process their offers, which constitutes a one-time setup that is independent of the number of buyer-triggered purchase inquiries. In contrast, in HPI and cHPI, the buyer has the most workload when preparing her (reusable) query. As identifying the "best-fitting" seller by contacting several sellers is a common application, this reuse is highly beneficial.

*RAM Usage.* Our designs do not require excessive amounts of memory at any time of the operation and are runnable on commodity hardware. Still, the memory usage

and its distribution among the participants differ. In PPI, the memory usage is driven by the Bloom filter as part of the PSI-based matching step. Consequently, it barely requires memory at the buyer (4 MB for 1 Mio. products) or the cloud (only around 200 B per matched product), but mostly burdens the seller (with a feasible RAM usage of 1 GB during the pre-processing of a product catalog 1 Mio. products). In comparison, HPI distributes the memory usage over both parties, seller and buyer (at most 1.9 GB for 1 Mio. products; with batching support). By design, cHPI relieves the seller of excessive needs because costly computations are offloaded to the cloud. Apart from that, both designs, HPI and cHPI, could be adapted to process the computation in batches, thereby lowering the maximum memory usage.

*Outgoing Transmissions.* The measured transmission needs are well feasible with today's infrastructures. While, in PPI, the seller has to transmit most of the data (only at most 4 MB, even for 1 Mio. products due to the efficient underlying Bloom filter), HPI and cHPI feature larger transmissions from the buyer to seller and cloud, respectively. For example, with $|X|$=5000 products, we measure around 9 MB of data transmissions. This number increases to 1.85 GB for $|X|$=1 Mio. Moreover, in cHPI, the seller needs to upload her blinded prices to the cloud, introducing notable overhead for her (in comparison to HPI). Thus, network limitations affect HPI's runtime, and especially cHPI's runtime, to a larger extent (in comparison to PPI).

Overall, we notice that HPI and cHPI require slightly more resources due to the PHE-induced overhead in comparison to PPI. Later, in Section 4.2.2.6, we discuss the privacy benefits of HPI and cHPI, which warrant the overhead for settings with the need for strict confidentiality guarantees. Regardless, we conclude that the performance is suitable for real-world deployments as our protocols can run on commodity hardware in a reasonable time, even when working with large product catalogs.

#### 4.2.2.5 A Real-World Setting for Purchase Inquiries

To verify the real-world feasibility of our designs, we studied two real-world applications, i.e., we relied on products from the domain of machine tools. In the following, we first introduce our considered datasets before discussing the evaluation results.

##### Queries in the Domain of Machine Tools

For our real-world evaluation, we consider two distinct queries on machine tools with an application in injection molding (cf. Section 3.2.2): Our queries exemplarily describe properties (of products) and offers from sellers of (a) clamping units where the material during the process is injected, i.e., the units determine the shape of produced workpieces, which we refer to as *tool query*, and (b) machines to produce such clamping units, which we label as *machine tool query* in the following.

A fundamental requirement for purchase inquiries is a unique product modeling that allows both potential sellers and interested buyers to have an unambiguous understanding of the product. Typically, a domain-specific set of parameters is used to define the modeling function. For an illustration of the different parameters, we

refer to our previous paper [PDF+23, Appendix C]. Exemplarily, we now present a subset of the relevant parameters in the *tool query*: (i) The *size factor* describes the maximum mounting length, which is crucial as machines are frequently limited in their workspace. (ii) The *shape complexity factor*, the *aspect ratio*, and the *filigree factor* identify the shape of the workpiece and, therefore, also of the mold. These aspects often introduce constraints on the tooling of supplier's machinery and thus reduce the number of suitable suppliers. (iii) *Tolerance* and *material factor* label the requirements concerning the surface quality of the product in question.

To combine the selected parameters into a modeling function that also unites value ranges of specific parameters (keeping the product catalog small), we define a binning scheme for each relevant parameter. We appropriately configure the respective parameter ranges and the bins' granularity for the respective application. For example, when querying for a tool, for the important size factor, we work with 5 bins, each covering a distinct set of tool configurations. In contrast, other parameters are only binary, e.g., indicating whether they feature dielectric operation. Finally, we source this binning scheme to discretize the products into product *id*s.

With this approach, we end up with 38 880 possible configurations (based on 10 parameters, each with 2 to 5 bins) in our *tool query* and 944 784 unique product configurations in our *machine tool query*. In the latter case, we express the different products of the product catalog through 14 parameters.

### Real-World Performance Measurements

With these applications, we are able to underline the real-world applicability of our approaches. The products in the catalogs (38 880 and 944 784 modeled products) match the values that we considered in our synthetic performance evaluation (cf. Section 4.2.2.4). As our designs are oblivious to the exact numbers that they are comparing, our measurements are in line with previous results.

*Tool Query.* For the first application, we report a total runtime of $2.3\,\text{min} \pm 0.7\,\text{s}$ for PPI and $21.6\,\text{min} \pm 0.2\,\text{min}$ for HPI, respectively. For cHPI, we measure a runtime of $21.6\,\text{min} \pm 5.2\,\text{s}$. Hence, offloading the computation of small real-world datasets does not influence the overall performance significantly with unconstrained network links. Again, PPI and our HE-based designs differ by one order of magnitude in runtime. In real-world settings with constrained network links, the speedup of PPI will be even higher due to the smaller total transmission size. Regardless, this application underlines that the designs allow buyers to quickly query sellers for available tools.

*Machine Tool Query.* Our second application features a significantly-larger product catalog. The overall runtime exhibits this aspect as well. PPI takes $57\,\text{min} \pm 15.5\,\text{s}$ to conclude its run, while HPI finishes after $525\,\text{min} \pm 3\,\text{min}$. With larger product catalogs, the offloading and additional blinding in cHPI slightly prolong the overall runtime of the design variant in comparison to HPI. Accordingly, we measure a runtime of $530\,\text{min} \pm 4\,\text{min}$ for cHPI. Given that the purchase of a new, complex machine tool (choosing from more than 940k product variants) is not an everyday query, which is usually planned well in advance, the real-world performance of our

designs is feasible for industrial settings. Thus, we argue that we could easily support even larger scenarios without limiting the value of our proposed purchase inquiries.

Conducting purchase inquiries is a critical task for many companies as they cannot produce every required part or production resource themselves. Thus, they are forced to trade. The presented performance of real-world applications underlines that identifying a suitable supplier from the multitude of possible suppliers on a global market is possible, even privacy-preservingly. Hence, in comparison to today's practices, buyers can easily consider a larger set of suppliers without risking the disclosure of sensitive information. In subsequent steps of the procurement process, companies can then also take other important factors, such as delivery times, into account. Overall, our designs protect sensitive information from unsuitable suppliers, i.e., effectively avoiding a competitive disadvantage for the inquiring company.

#### 4.2.2.6 Security Discussion

After studying the performance of our designs for two-way privacy in purchase inquiries, we now discuss their security (guarantees). In this regard, we primarily consider buyer and seller privacy (**G-I1** and **G-I2**). Moreover, we look at the protocol resistance (**G-I3**) of our designs to ensure their acceptance in real-world deployments.

As part of our security discussion, we consider malicious-but-cautious entities (cf. Section 2.1.2.1). Due to the authenticated communication during purchase inquiries, all participating companies are identifiable. In the following, we further detail that they do not have any incentive to input incorrect information into our protocols. We envision that a trusted industry association, such as the VDMA [VDM15], operates—funded by membership fees—the third party as a public service when using PPI or cHPI. Naturally, our work bases on the security of the established secure communication channels, the used (technical) building blocks, and properly-chosen key lengths (ensuring an adequately-secure mode of operation).

**Bilateral Protocols.** All designs concern a single buyer-seller pair only: For security, we utilize secure two-party building blocks from confidential computing (PSI or HE) in our protocols. Given that all protocol runs (i) are independent of each other and (ii) do not source third-party data, only the involved parties can potentially attack the protocols with the goal of extracting information. Uninvolved parties, especially other (uninvolved) buyers and sellers, cannot extract any information.

**Information Leakage.** Theoretic information leaks for participating parties are limited by design: First, a buyer could repeatedly send queries to a single seller to brute-force her prices or re-construct her offered product catalog. Given that no centralized third party is handling the purchase inquiries, sellers can freely decide whether they want to apply some kind of rate limiting for any buyer to enforce their privacy needs (**G-I2**). However, we assume that, in practice, no such action is needed as the workload for buyers renders frequent requests unlikely (cf. Section 4.2.2.4). Second, due to the weaker privacy guarantees in PPI (resulting from its two-phased design), a buyer could fear that her requested products are leaked (partially) during the comparison phase. However, we only work with granular capabilities, and

thus no, detailed product information is revealed at any time. Next, we individually discuss the respective implications of each design.

*PPI: A Two-Phased Approach.* We generally ensure buyer and seller privacy (**G-I1** and **G-I2**) in both phases. First, the matching step relies on PSI and only returns the matches to the buyer, i.e., the seller cannot learn anything from this phase (ensuring **G-I1**). In theory, the buyer could request every possible product to derive the seller's catalog. However, as the seller is involved in the preparation of the buyer's query, the seller can rate-limit the buyer if she seems to be requesting too many products. For the same reason, buyers cannot brute-force products listed in Bloom filters that index the seller's product catalog. Thus, apart from sharing intermediate results on matches with the buyer, we also address seller privacy (**G-I2**) in the matching step.

In the comparison step of PPI, we rely on a third party (cloud). The used ORE ciphertexts are never shared with any other party except for the cloud, which only operates on these ciphertexts. Thus, as intended (also a design goal of ORE), it can only learn the ordering of any two ciphertexts without any knowledge of the compared products. As buyers and sellers agree on a shared key $k$, the cloud cannot analyze comparisons over time, limiting the cloud's insights to a single query, i.e., the cloud cannot perform frequency analysis to, e.g., uncover frequently-requested products. However, PPI can "leak" the matches (prior to the price comparison) to the seller as the cloud only queries the prices for requested matches by default. At the expense of moderate performance overhead, e.g., a runtime of 100 s to encrypt 1 Mio. prices, we could require the seller to share ORE ciphertexts for all prices by default, effectively preventing said intermediate information leaks. Thus, in PPI, we can tune the fulfillment of **G-I1** and **G-I2** during the comparison.

*HPI: An HE-Based Protocol.* HPI is secure and ensures privacy by design [FG07]: The buyer encrypts its data homomorphically (**G-I1**), and the contacted seller never provides any information to others. Thus, no data, except for the final result, is shared as all computations operate on PHE ciphertexts. The seller can, at no point, decrypt the buyer's data or the result. To ensure confidentiality (**G-I2**), the seller blinds the result before returning it to the buyer (who learns a single result).

*cHPI: A Cloud-Tailored Design Variant.* Like HPI, cHPI is secure and preserves the confidentiality of sensitive information by design as it still bases on PHE, i.e., only the buyer can decrypt the ciphertexts. All remaining parties (seller and cloud) directly operate on ciphertexts, i.e., they only compute data homomorphically.

Since we consider the third party (cloud) in cHPI to be untrusted, it must not have access to any data in the clear or observe any insightful patterns. To ensure privacy while offloading unencrypted inputs to the cloud, the seller must blind her price expectations to still achieve seller privacy (**G-I2**). To this end, buyer and seller jointly agree on a random vector $N$. Consequently, the cloud cannot conduct any correlation attacks as the individual prices are obfuscated in a different way for each protocol run and query. The buyer's pre-processing capabilities are not limited in any way by this step as the cloud still receives the unblinded PHE ciphertexts $C_B$. Using the scalar product $\langle N \cdot C_B \rangle$, the seller can eventually remove the random vector $N$, i.e., she locally unblinds the cloud-computed result. As for HPI, the seller

also blinds the result in cHPI to ensure confidentiality before returning the result to the buyer. Thus, we still establish seller privacy (**G-I2**) when deploying cHPI.

**Entity Misbehavior.** Buyer and seller could still behave according to the protocol while providing manipulated queries or product catalogs, respectively. First, in theory, the buyer could send bogus queries. In this case, the protocols return matches on products that the buyer is not interested in, i.e., the information has no added value to him. In addition to his own resources, such misbehavior would, however, also consume computational resources of the seller. Second, the seller could compile a fake catalog to increase the probability of matches and, thereby, the likelihood of subsequent negotiations. By inserting additional products, the seller might know what products the buyer is interested in (if being contacted). However, as he cannot produce them, no sale will be made (cf. **G-I3**), and his reputation will suffer. Higher prices would even lower the probability, and getting the buyer to negotiate by specifying lower prices either leads to lower prices during the sale, i.e., no misbehavior, or no sale at all (cf. **G-I3**). Thus, entity misbehavior leads to no benefits.

**Collusion Attacks.** Possible collusion among multiple parties of a bilateral purchase inquiry could potentially increase the illegitimate information gain for misbehaving parties. Thus, we now discuss the threat of multiple colluding parties. Due to the lack of a third party, such attacks are not possible in HPI.

*Buyer and Seller.* Such collusions contradict our setting (cf. Section 4.2.1.1) as these parties can exchange all sensitive information directly anyway. Besides, they cannot gain any additional (third-party) data as our protocols are bilateral by design. Thus, such collusions are irrelevant to assess the security of our designs.

*Buyer and Third Party.* In PPI, these parties can reveal the prices $min_{id}^{S_i}$ of (matched) products during the comparison step due to the buyer's knowledge of the shared key $k$, i.e., he knows Seller $S_i$'s lowest selling price. Thus, the seller's privacy (**G-I2**) depends on a carefully-selected third party. Similarly, if a buyer is colluding with the third party when using cHPI, they can jointly extract all price expectations of the seller, e.g., to only offer the seller's lowest selling prices in subsequent negotiations: The third party has access to the blinded prices, and the buyer knows the random vector $N$, which can be used to remove the applied blinds. Hence, cHPI cannot tolerate such a collusion to ensure seller privacy (**G-I2**). However, when relying on a trusted operator, we can effectively reduce the probability of such an attack.

*Seller and Third Party.* In PPI, such a collusion can reveal the prices $max_{id}^{B}$ of matched products during the comparison step due to the seller's knowledge of $k$, i.e., he is aware of Buyer $B$'s highest purchase price. Thus, a carefully-selected trusted third party is recommended to ensure buyer privacy (**G-I1**). A collusion of the seller and the third party in cHPI has no negative consequences for the buyer because both parties operate with PHE ciphertexts anyway; and without the encryption key, they cannot decrypt the ciphertexts. Thus, such collusions have no effect on buyer privacy. If the seller trusts the third party, we could even refrain from blinding the price expectation using $N$, reducing the overhead of the cloud offloading in cHPI.

To conclude, due to the shared key $k$ in PPI, we cannot tolerate any collusion with the third party as it would allow for leaks of the minimum, respectively maximum

prices, slightly violating **G-I1** and **G-I2**. Thus, for settings with strict privacy needs, HPI or cHPI should be used instead. Regardless, we consider PPI to be resistant in terms of **G-I3** as such collusions do not influence the outcome of PPI's computation.

**Multiple Buyers.** Such a collusion contradicts our defined attacker model as syndicated procurements [OECD13a] are only permitted through a wholesaler, which in turn equals a single buyer in our protocol use, i.e., our protocols are not affected by any (buyer) collusion in this case.

**Multiple Sellers.** Just like a collusion of multiple buyers, a collusion of multiple sellers contradicts our defined attacker model as cartels [OECD13b] are forbidden by law. Consequently, they cannot legally use our protocols in practice to, e.g., globally drive up product prices. Hence, such (seller) collusion is out of scope.

Our security discussion underlines the privacy benefits of HPI. Likewise, cHPI allows companies to realize secure (and privacy-preserving) purchase inquiries while simultaneously reducing the required resources for participating sellers. Thus, it serves as a suitable alternative for situations where (local) computing resources are scarce, despite its disadvantages of networking overhead and additional operating costs. However, even with the intuitive and more performant PPI, most information is kept private by design as part of our purchase inquiries.

#### 4.2.2.7 The Potential of Private Purchase Inquiries

Our evaluation underlines that our designs are suitable for our considered scenario (cf. Section 4.2.1.1). With our HE-based protocols, HPI and cHPI, we are further able to address all design goals (cf. Section 4.2.1.2) to enable privacy-preserving purchase inquiries as part of the established procurement process in the industry.

While our protocols generally ensure buyer and seller privacy (**G-I1** and **G-I2**; PPI with minor deductions) by design (cf. Section 4.2.2.1), we demonstrate their real-world feasibility (**G-I4**) using our machine tool applications (cf. Section 4.2.2.5). These runs conclude within an appropriate time and reasonably consume resources to be suitable for real-world use. During our security discussion (cf. Section 4.2.2.6), we further looked at our protocols' robustness (**G-I3**): Overall, we present secure protocols as their security mainly builds on established building blocks with attested security.

When considering both performance and security, we notice the trade-off between no information leakage in HPI and cHPI at the expense of some computational overhead when compared to PPI. More specifically, in settings where a two-phased protocol with an ORE-based price comparison, which is conducted by a cloud service, is acceptable, companies can benefit from the superior performance of PPI. Its performance is superior to HPI (and cHPI) because the protocol only involves the seller's catalog $P_c^{S_i}$ and not every product $id$. Thus, in practice, fewer product $id$s are part of the pre-processing and the comparison. Apart from fewer computational resources, PPI also supports indicating matches for each queried product, i.e., queries are answered with more granularity. However, adjusting the cloud's

protocol to only return a single result, as in HPI (cf. our introduced notation in Section 4.2.2.1), is a simple, non-invasive adjustment. Thus, according to their individual needs, companies can select a design with technical guarantees that provides satisfactory and performant two-way privacy during procurement.

Our approaches further scale well with the number of potential sellers that should be considered as the protocols runs are independent of each other, and buyers can easily trigger as many purchase inquiries as computationally supported. Thus, due to this flexibility, our proposed designs also fulfill the targeted goal of ease of use (**G-I5**).

## 4.2.3 Takeaways and Future Research

In this section, we have introduced our designs for two-way privacy as part of purchase inquiries during procurement. With this contribution, we account for the trend toward highly-dynamic supplier networks in the IIoT, which entails that the number of business relationships among mutually-untrusted stakeholders increases. In particular, we provide companies with protocols that protect their sensitive information during these early stages of procurement through technical building blocks. We now conclude our presentation by first discussing the suitability of our selected building blocks in Section 4.2.3.1. Afterward, in Section 4.2.3.2, we briefly highlight the conceptual implications of our privacy-preserving purchase inquiries on traditional procurement processes. Finally, in Section 4.1.4.3, we outline potential future work.

### 4.2.3.1 Suitability of Selected Technical Building Blocks

Based on the derived design goals (Section 4.2.1.2), primarily **G-I3** to **G-I5**, we follow that the range of potential building blocks to choose from is already constrained. The goal to bilaterally protect purchase inquiries, i.e., implementing independent protocol runs, hinders the application of sophisticated secure multi-party computations. Thus, we resort to "simpler" designs using building blocks from the area of privacy-preserving computation, such as HE and PSI. The application of the latter is restricted by its ability to "match" two dimensions (i.e., product(s) and price range(s)) simultaneously. Thus, for PPI, we had to rely on a two-phased design, which results in slightly-weaker confidentiality guarantees (due to the third-party-based ORE comparison), as we have outlined in Section 4.2.2.6. In contrast, sourcing HE to realize two-way privacy appears to be a straightforward candidate with strong security guarantees if a suitable algorithm to conduct the comparison can be found because the performance goal of concluding a purchase inquiry within several hours (cf. **G-I4**) leaves room for demanding approaches. With HPI and cHPI, we even derived underlying computations that allow us to resort to PHE, which reduces ciphertext sizes and the computational overhead of applying HE (when compared to FHE). Our choice to study multiple designs that build on different building blocks further allowed us to outline the performance and security implications of these concepts. Overall, we are confident that our chosen building blocks are appropriate to realize practical privacy-preserving purchase inquiries for the first time.

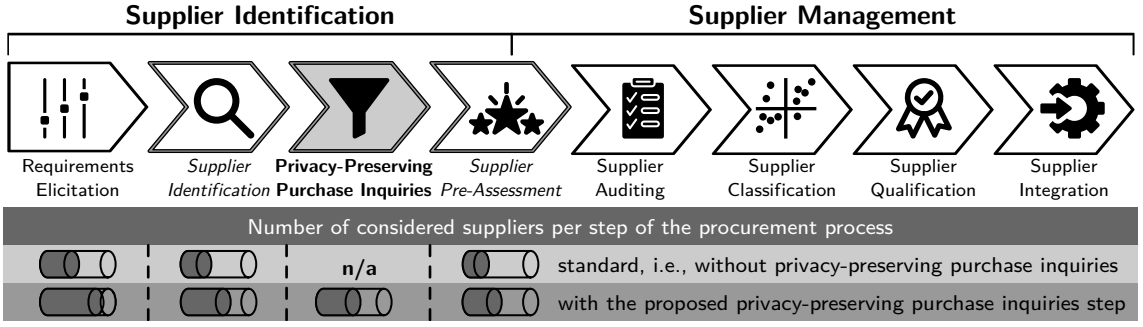



**Figure 4.18** We augment existing, established procurement processes [BKdL+18] with a step in the supplier identification. Thereby, we can increase the set of considered suppliers in practice.

#### 4.2.3.2 Implications for Purchase Inquiries

Procurement in business-to-business markets is a complex task with several steps, as we illustrate in Figure 4.18. For a detailed description of these steps, we (again) refer to our previous paper [PDF+23, Appendix C]. In the following, we instead focus on the practical impact of conducting purchase inquiries privacy-preservingly.

When considering the sharing and revelation of (sensitive) information, the most critical steps are at an early phase of the procurement process, mainly because it also involves completely unknown, likely untrusted suppliers (sellers). Thus, buyers cannot forecast how these potential sellers will process and handle their information, i.e., they fear for their competitive advantage. Likewise, sellers are also not interested in disclosing all of their capabilities, as they also fear damaging effects. To account for this dilemma, we propose the introduction of a new intermediate step—privacy-preserving purchase inquiries (cf. Figure 4.18)—that directly integrates into existing procurement processes. Due to its non-invasive nature, it only slightly affects the traditional steps (*supplier identification* and *supplier pre-assessment*) as parts of these steps are now covered by our newly-added step. All remaining steps of traditional procurement processes remain completely oblivious to this change, i.e., companies can, for the most part, stick to their established practices.

Our new privacy-preserving purchase inquiries step is highly beneficial for use in real-world deployments in the IIoT, because it allows buyers to consider significantly more sellers at the early phases of the procurement process (supplier identification phase). Using technical means, we ensure that no sensitive information is leaked to other parties and especially unsuitable sellers. Given the larger number of initially-considered sellers (visualized at the bottom of Figure 4.18), buyers might be able to source their sellers from a large(r) set of fitting suppliers. With today's practices, these suppliers might not have been considered at all (cf. Section 4.2.1.1).

#### 4.2.3.3 Future Work and Next Steps

To the best of our knowledge, our work is the first in the area to address confidentiality issues during procurement. Consequently, we expect that research efforts in this direction will significantly increase to holistically address this overlooked area.

Research could build on our initial work to, for example, look into the possibility of improving PPI by replacing our ORE-based price comparison with another secure approach, e.g., secure multi-party computation or homomorphic encryption, to address the slightly-impaired privacy guarantees in PPI. Despite this room for improvement, our work details how to offer two-way privacy for buyers and sellers while outlining how to neatly integrate our designs as an immediate next step of procurement. Detached from this purely-technical view, we identify some follow-up research questions that would help to resolve potential reservations against the acceptance of our novel privacy-preserving purchase inquiries. As a simpler direction, related work should study the configuration of rate-limiting approaches to prevent (i) repetitive, possibly-iterative purchase inquiries with a single seller or (ii) denial-of-service attacks on sellers or groups of sellers. A more challenging aspect concerns the *external* verifiability of conducted purchase inquiries, e.g., to prove to the government that a number of offers have been requested. So far, we see a lack of real-world deployable solutions that protect all confidential inputs while also allowing (external) parties to verify the conducted computation. Even though our work on improving the finding and bootstrapping of new suppliers for business relationships along the supply chain as part of the procurement process would benefit from corresponding building blocks, we consider this challenge to be an open and essential research direction for the general area of privacy-preserving computation. Finally, to accurately study the real-world implications of our work, we call for economic studies measuring the (monetary) impact of our privacy-preserving purchase inquiries in industry. Likewise, we look forward to real-world acceptance studies to confirm the attested ease of use of our designs. We are confident that once the evolution of this crucial aspect picks up, further refined solutions, which improve the maturity of privacy-preserving purchase inquiries in the IIoT, will emerge both in academia and industry.

This subsection concludes the presentation of our second contribution, which greatly supports the risk-free establishment of dynamic buyer-seller relationships along supply chains. Our presented designs improve the status quo in today's manual, privacy-invasive procurement by offering two-way privacy for the involved companies. With our practical, real-world feasible designs and our corresponding discussions, we make an important step toward a secure, productive, and efficient industrial landscape: We provide industry with means to utilize secure collaborations even in the context of mutually-distrusting stakeholders. Overall, in this chapter, we have advanced two types of collaborations along the supply chain. First, by enabling multi-hop and privacy-preserving information flows that follow from established business relations (Section 4.1), a second, by addressing the confidentiality concerns of companies at the early stages of the procurement process. While we consider the expected benefits and implications (of our work and collaborations along supply chains) for an evolving IIoT to be significant, in the next chapter, we shift our focus to (secure) collaborations across supply chains, an aspect that has rarely been studied so far.

# 5

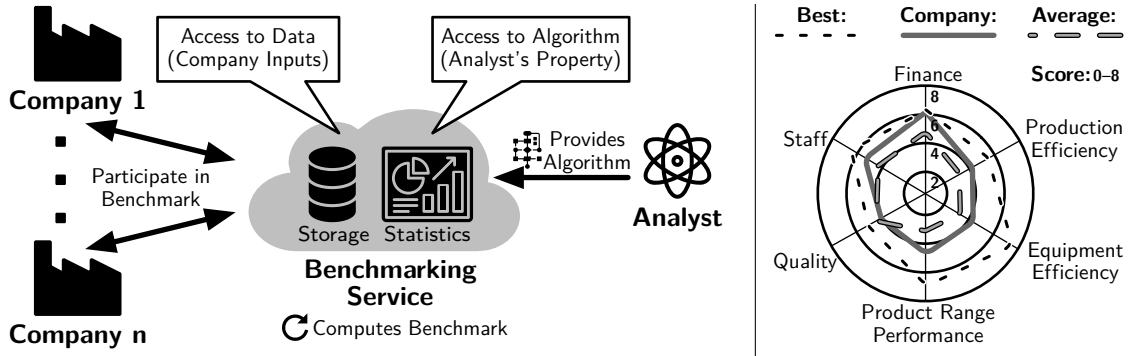
# Collaborations Across Supply Chains

In this chapter, we look at (secure) industrial collaborations across supply chains. As in the previous chapter, we again focus on two different settings. That is, for our third contribution, in Section 5.1, we first study how to privacy-preservingly realize real-world company benchmarks, primarily in settings with established business relations. In this context, we specifically consider the sensitivity of the underlying benchmarking algorithm, which has largely been overlooked in the past. Afterward, for our fourth contribution, we shift to a setting with unknown, most likely untrusted collaborators. In Section 5.2, we detail how to establish secure collaborations that allow companies to privacy-preservingly exchange information across supply chains (and domains), a feature that companies in the IIoT await eagerly. To address this desire, we present an architecture that is oblivious of the processed information.

As we have discussed in Section 1.3, collaborations across supply chains differ from collaborations along supply chains as (i) privacy concerns are prevalent and (ii) information flows are not yet widely established. With these challenges in mind, we first look at a secure collaboration without direct implications on running processes (the least invasive type of collaborations across supply chains). Specifically, our third contribution covers privacy-preserving comparisons. In contrast, our fourth contribution on privacy-preserving matchings entails more precarious consequences as the retrieved/exchanged information is likely fed directly into running processes.

## 5.1 Privacy-Preserving Company Benchmarking

Due to the fundamental pursuit of monetary interests, businesses always try to improve their processes (and products) while targeting a variety of goals (cf. Figure 1.1). As a foundation for any changes, companies need to *reliably* identify potentials for improvement. In this context, company benchmarks are a common industry practice




**(a)** Benchmarks rely on sensitive inputs and an algorithm **(b)** Visualization of KPIs

**Figure 5.1** Using the analyst's algorithm and the companies' inputs, the benchmarking service computes all target KPIs, allowing companies to accurately assess their performance.

to reveal shortcomings with respect to business partners and competitors alike, i.e., corresponding comparisons are challenged by the sensitivity of sourced information.

In the following, in Section 5.1.1, we first introduce company benchmarking along with its prevalent privacy challenges for the involved stakeholders. Afterward, in Section 5.1.2, we present and evaluate our designs that address these privacy challenges while providing technical guarantees. Finally, we conclude the presentation of our third contribution on privacy-preserving business comparisons in Section 5.1.3.

## 5.1.1 Privacy Issues in Company Benchmarking

As a foundation for our third contribution, in Section 5.1.1.1, we give a brief recap of company benchmarking (cf. Section 3.2.3) and outline its impact on businesses in the IIoT. Subsequently, in Section 5.1.1.2, we detail the algorithms that are being sourced in real-world benchmarks. Based on these examples, we then derive general design goals (Section 5.1.1.3) that capture the privacy challenges of company benchmarks. Afterward, in Section 5.1.1.4, we discuss related work and highlight the lack of approaches that also consider the confidentiality of the underlying benchmarking algorithms. Moreover, we additionally present relevant privacy-enhancing concepts (proxy re-encryption and k-anonymity) in Section 5.1.1.5. Based on this foundation, we then detail our designs that preserve the privacy of all involved stakeholders.

### 5.1.1.1 Company Benchmarking in Industry

Company benchmarks usually focus on practices such as the company's operations and the management of a company or a department [MdRC12]. The main objectives are to evaluate the company's current market position in relation to a recognized leader (or certain peer groups [Ker08]), as well as to adapt local processes to close any gaps, e.g., by avoiding a waste of resources [Tei01, Koz04]. It provides companies with insights into the effectiveness of their current processes (qualitatively and quantitatively). For example, Xerox, a manufacturer of photocopiers and document management systems, improved its annual productivity gains from 3 % to 5 % to

| Dataset | Inputs | KPIs | Max. Depth | Avg. Depth | Formulas | Operations |
|---|---|---|---|---|---|---|
| IM | 674 | 48 | 49 | 12 | 627 | 2704 |
| PN | 35 (n-dim) | 14 | 12 | 6 | 14 | 100 |

**Table 5.1** Overview of our real-world applications and their respective algorithm complexity.

10 % [TZC87] after comparing its processes with L.L. Bean, a retailer of outdoor sporting goods, and addressing the benchmark's findings.

Benchmarks operate on key performance indicators (KPIs), allowing for quantitative comparisons of products, services, or implemented practices [Ker08,HSF+09]. Nowadays, the sets of relevant KPIs frequently change. For instance, they increasingly cover sustainability, which also allows for comparisons of environmental and social aspects [Boo21]. In Figure 5.1a, we illustrate the process of benchmarking, including the main actors: an analyst, the benchmarking service, and participating companies. First, the analyst develops suitable algorithms to compute meaningful KPIs, which are usually kept private due to their value and intellectual property [GPSPD06]. A relatable example, which involves consumers, are credit scoring agencies, such as Experian or Schufa, which largely depend on the confidentiality of their algorithms. The benchmarking service collects the required inputs from participants (companies in our setting) and computes the KPIs to compare them as part of the benchmark. Eventually, the participants receive the general results and their own KPIs. By studying the results, participating companies can then investigate their performance in comparison to the average and "best in class", as we illustrate in Figure 5.1b.

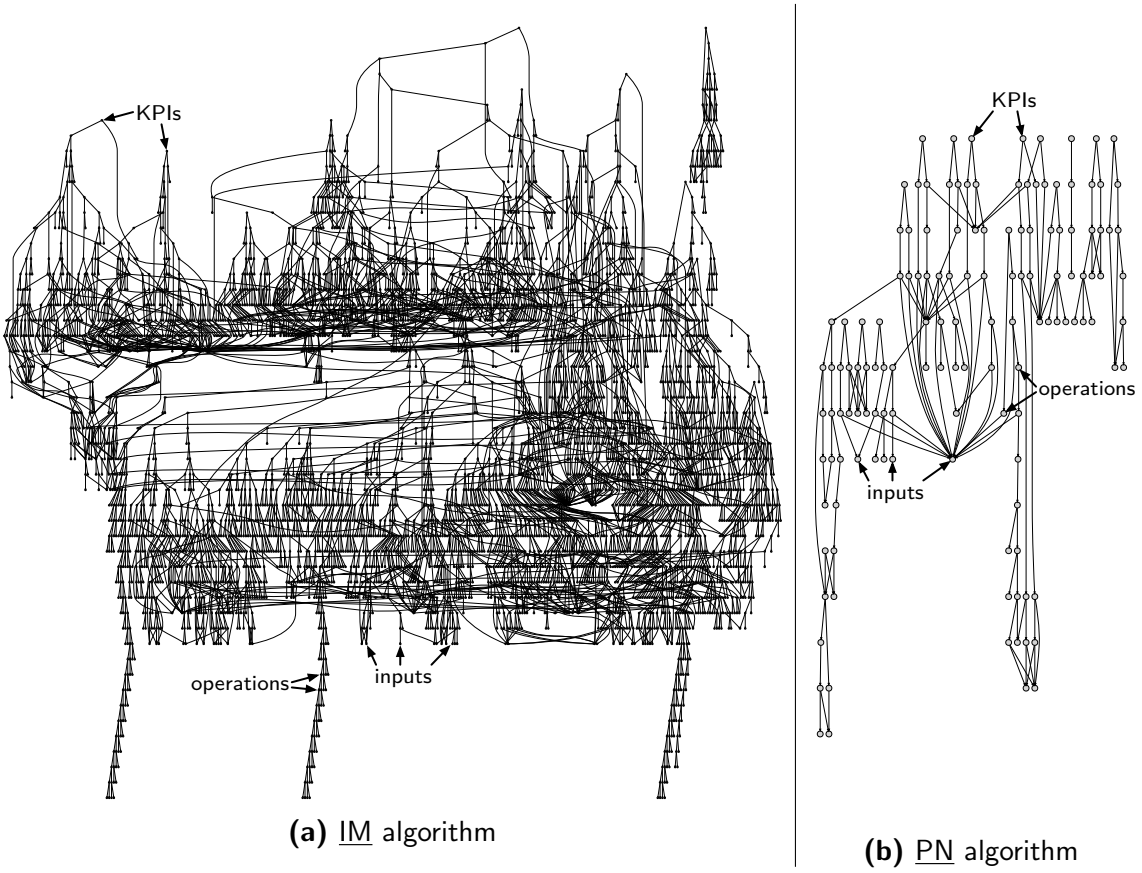
#### 5.1.1.2 Real-World Applications of Company Benchmarking

In the following, we pick up our two real-world applications from Section 3.2.3 to highlight the specifics of real-world benchmarks in more detail. Afterward, as a foundation for deriving the design goals, we explicitly stress the complexity of the underlying algorithms that are central to the respective benchmark.

##### Benchmarking Companies in Injection Molding (IM)

Our first benchmarking application covers injection molding. Accordingly, the underlying algorithm and hence most of the resulting KPIs are highly specialized for this domain. As we detail in Table 5.1, the complexity of this example is high, with computations in up to 49 sequential operations, over 600 inputs, and more than 2700 operations, i.e., the analyst's effort of crafting KPIs is significant. In addition to elementary arithmetic, the KPI computation also sources exponentiation ($x^y$), roots ($\sqrt[n]{x}$), as well as absolute ($|x|$) and extrema values (min/max). In return, participants receive detailed results due to the large number of KPIs (Figure 3.2b highlights some of them). As such, this use case is a representative real-world example, and it is along the lines of the number of expected KPIs (cf. Section 5.1.1.4).

For an in-depth presentation of the setup and the company benchmark itself within this application, we refer to our previous paper [PSF+20, Section 5.2].




**(a)** IM algorithm **(b)** PN algorithm

**Figure 5.2** The complexity of our real-world benchmarks differs by orders of magnitude. Computing a single KPI usually requires several inputs and depends on complex layered formulas.

### Measuring the Efficiency of Global Production Networks (PN)

Our second application benchmarks the performance of production sites in globalized production networks. In comparison to IM, the underlying algorithm of PN features three interesting differences: (i) arrays as input values with variable length, which might implicitly reveal sensitive company details, (ii) component-wise operations on arrays, and (iii) summation ($\Sigma$) or extrema over arrays. Hence, despite its small size, with 14 KPIs and 100 operations (cf. Table 5.1), it is of great relevance when designing benchmarking systems due to the complex operations contained within.

### The Complexity of Real-World Benchmarks

These real-world benchmarks stress the importance of ensuring the confidentiality of benchmarking algorithms in real-world deployments because the derivation of meaningful KPIs is a labor-intensive activity. To the best of our knowledge, related work usually neglects this effort. The presented benchmarks show that KPI computations can be very complex, i.e., a single KPI can be based on several formulas with dependencies, diverse operations, and hundreds of inputs. Moreover, a single benchmark may even consist of up to 200 KPIs [Ker08]. In Figure 5.2, we exemplarily illustrate the layered structure and deeply-rooted formula dependencies for both applications.

#### 5.1.1.3 Design Goals for Practical yet Privacy-Preserving Benchmarks

Related work [Ker08, SMF18, BBS19] frequently outlined the need for confidential KPIs in the past. However, the prevalence of closed benchmarking algorithms stresses the need to also protect the computation of the compared KPIs as it represents the analyst's competitive advantage who invests significant effort to derive meaningful KPIs [PSF+20]. In this context, we identify three crucial dimensions when securing company benchmarks: (i) *benchmarking frequency* (one-time vs. continuous benchmarking), (ii) *openness of data* (i.e., open vs. closed data), and (iii) *openness of the algorithm* (i.e., open vs. closed benchmarking algorithms).

Traditionally, benchmarks utilize labor-intensive manual interviews [Koz04, PSF+20] to collect the data that subsequently feeds the KPI computation, i.e., we can classify them as one-time benchmarks that source closed data and a closed algorithm. Data-driven approaches increasingly evolve toward continuous benchmarks, which might additionally rely on public (open) data and algorithms (e.g., governmental applications). As such benchmarks (open data and open algorithms) do not require elaborate security mechanisms, they are comparably easy to realize. Contrary, company benchmarks require strong security as they operate on sensitive (closed) company inputs using valuable, use case-tailored, and complex (closed) algorithms.

With this background in mind, we now derive a set of five distinct, general design goals to truly enable privacy-preserving company benchmarks in the IIoT.

**G-B1: Company Privacy.** Both the raw inputs required to calculate KPIs and individual KPIs of companies have to be processed with care as they could reveal critical information to competitors. Well aware of these risks, companies nowadays are extremely reluctant to participate in centralized benchmarking systems that require access to data in plain text [Hen20]. Hence, when fully considering these confidentiality needs, companies are likely to increasingly participate in benchmarks.

**G-B2: Complexity.** Contrary to the common misconception of related work, real-world benchmarks frequently build on complex and often hierarchical formulas to compute meaningful KPIs. In practice, simpler KPIs could potentially prevent meaningful comparisons among different companies [Par15]. Our two real-world applications (cf. Section 5.1.1.2) further underline the aspect of algorithm complexity.

**G-B3: Algorithm Confidentiality.** As we have presented in Section 5.1.1.2, the derivation of impactful and commercially-attractive benchmarks is a costly and time-consuming process [Par15]. Even for KPIs with seemingly simple calculations, significant upfront effort by the analyst might be required to compose them in a meaningful way. Consequently, these algorithms should be treated as sensitive information as they are the analysts' intellectual property and competitive advantage.

**G-B4: Exactness.** Since KPIs can build on complex hierarchical computations, with comparison results possibly influencing business decisions, ensuring the correctness of the performed calculations is essential. Accordingly, this requirement forbids distorting or abstracting values intentionally to protect the participants' privacy.

**G-B5: Flexibility and Scalability.** The participation of as many companies as possible is desirable to reach the full potential of company benchmarking [ABL+04]. Consequently, corresponding benchmarks should be easy to use for participating companies, i.e., require only a limited setup and no explicitly-trained staff. Likewise, participants should need to upload their contributed values only once, without the requirement to remain available long-term. Finally, to provide long-lasting flexibility, algorithms should be updatable, including the possibility to introduce entirely new functional building blocks, e.g., new mathematical operators. Likewise, company benchmarks need to scale independently of the number of participants as the usefulness of benchmarks increases with every new participant [ABL+04], making it pivotal to easily scale with the number of benchmarked companies in a single setup.

These design goals express several crucial aspects. First, after addressing the confidentiality needs of companies (**G-B1**), an increasing number of participants will also increase the usefulness of the benchmark due to its broader data basis, which will also generate more revenue for the benchmarking service (and analyst). Second, and on top of an increase in revenue, sufficiently-protected algorithms (**G-B3**) would counteract potential losses of subsequent compensations through unauthorized and unpaid reuse of algorithms. Hence, secure designs could additionally persuade analysts to invest resources in deriving valuable algorithms and KPIs. Thus, overall, real-world-applicable designs should carefully address these goals to ensure deployability and usability even for large-scale practical scenarios while allowing as many participating companies as possible to benefit from privacy-preserving benchmarks.

#### 5.1.1.4 Related Work

Traditional benchmarking services frequently utilize centralized designs that digitize paper-based responses of participants before computing KPIs and comparing them [PSF+20]. Apart from their labor-intensive realization, such benchmarking services conceptually serve as a trusted third party as they have access to all sensitive inputs. Such centralized designs protect the algorithm but fail to account for the sensitive company inputs (**G-B1**), hindering real-world adoption and the willingness of companies to participate [FPR04]. In contrast, local computations by the participants, who only return the computed KPIs, protect sensitive inputs but fail to account for the required algorithm confidentiality (**G-B3**). In a general direction, advances in privacy-preserving data processing emerge in research [vdA21]. However, they frequently build upon disclosing the utilized algorithms as well.

For a detailed comparison of related work, separated into approaches with (i) client computation, (ii) a trusted third party, (iii) secure multi-party computation, and (iv) concepts utilizing multiple servers, we refer to our previous paper [PSF+20]. Here, we focus on ongoing developments in the areas of *privacy-preserving computation* and *confidential computing* and summarize prior benchmarking designs.

**Software-Based Approaches.** In related work, we discover several software-based designs utilizing secure multi-party computation or homomorphic encryption. The former approaches usually have two major drawbacks: (a) they are commonly

round-based, i.e., all participants need to participate simultaneously [Ker08,BBS19], and (b) the scalability is, at best, quadratic [Ker11,BBS19] in the number of participants. Hence, these designs contradict **G-B5**. Initial HE-based approaches [SMF18, SKEEA18] come with a limited set of supported operations (violating **G-B2**) that challenge the computation of complex operations directly on encrypted data. All of these approaches do not consider the need for algorithm confidentiality (**G-B3**), i.e., they only protect the comparison of KPIs (**G-B1**) but fail to account for the sensitivity of the KPI computation (the analyst's intellectual property).

As a related research direction, prior work [Ker07, HSF+09, Ker11, SSK+13] studied the influence and composition of peer groups on the participants' privacy. We consider this line of research as orthogonal, and, in this dissertation, we focus on the privacy-preserving computation of KPIs without leaking the algorithm instead.

**Hardware-Based Approaches.** In contrast to privacy-preserving computation, confidential computing and corresponding hardware-based concepts only emerged after the *majority* of software-based benchmarking services had already been proposed. We were generally unable to discover benchmarking services that rely on TEEs. Irrespective of this focus, the range of applications that utilize TEEs to securely execute programs is immense, with them also moving toward mobile devices, such as smartphones, these days. For example, TEEs are in use to securely execute user applications [BPH15], improve the security of docker containers [ATG+16], protect voice assistants in the cloud [BFR+18], or run Tor in a shielded environment [KHH+18].

Detached from benchmarking services, the challenge of collecting data from different sources to compute statistics, comparisons, or benchmarks has been studied from different angles, mostly centering around differential privacy, secure multi-party computation, and homomorphic encryption. In settings that primarily involve private users, different approaches tackle the challenge of securely crowdsourcing statistics from user devices [BOT13, EPK14], perform statistical queries over distributed data [CRFG12, CAF13], or nudge users to more privacy-conscious behavior based on comparisons [ZHHW15]. All these approaches have in common that they focus on the confidentiality of user data using differential privacy to carefully distort aggregate statistics. While this focus is a reasonable trade-off when considering private users, company benchmarking involves complex and nested calculations of KPI (**G-B2**) and demands a high level of correctness (**G-B4**), contradicting the design goals of differential privacy, which mainly concentrates on hiding the data's origin.

**Research Gap.** While various conceptual approaches in the area of company benchmarking have been proposed, they all assume that KPIs are readily available for (privacy-preserving) comparisons, neglecting the process of deriving them. However, such algorithms are extremely valuable, and ensuring their confidentiality is, therefore, a key concern of the analyst. Unfortunately, related work fails to address this need by solely focusing on the participants' privacy (**G-B1**), disregarding **G-B3**.

Accordingly, in this dissertation, we holistically study the suitability and applicability of hardware-based and software-based concepts for secure benchmarking services.

#### 5.1.1.5 Preliminaries: Privacy-Enhancing Concepts

In addition to the building blocks that we have already presented in Section 2.3, we later source two additional privacy-enhancing concepts, namely, proxy re-encryption and k-anonymity, in our designs. Therefore, we introduce them in the following.

**Proxy Re-Encryption.** The concept of proxy re-encryption specifies the ability of a third party (the proxy) to re-encrypt ciphertexts for another recipient without the need to decrypt the ciphertext [AFGH06], i.e., informally speaking, the underlying encryption key is substituted with another encryption key. Given the lack of decryption capabilities, the proxy does not have to be a trusted third party.

**k-Anonymity.** Even if data is properly anonymized, i.e., all identifiers have been stripped, unique values or value combinations might still allow (external) entities to draw conclusions on the entity contributing these values [Swe02]. Likewise, side-channel information such as timing information [PH10], i.e., at which point in time data has been submitted, can aid in de-anonymizing entities. To prevent such inference attacks, the concept of $k$-anonymity [Swe02] suggests skillfully creating an anonymity set of size $k$. Hence, to apply k-anonymity to benchmarks, the KPIs of at least $k$ participants have to be aggregated before their (public) disclosure.

### 5.1.2 Designs for Privacy-Preserving Company Benchmarks

To extend related work with approaches that consider the need for confidentiality of both the company's sensitive data and the valuable algorithm, we propose two reference designs for **P**rivacy-preserving **C**ompany **B**enchmarking (PCB). Depending on their underlying building blocks, we refer to them as **H**ard**w**are- or **S**oft**w**are-based PCB, i.e., HW-PCB and SW-PCB. The latter design SW-PCB is an evolution of our initial PCB design, which we presented in detail in a previous paper [PSF+20].

In the following, we first provide a high-level overview of our designs' processing steps in Section 5.1.2.1. Subsequently, we detail HW-PCB and SW-PCB in Sections 5.1.2.2 and 5.1.2.3, respectively. In Section 5.1.2.4, we present our corresponding implementations and synthetic measurements of their performance. Afterward, in Section 5.1.2.5, we continue with our evaluation of two real-world applications, namely, IM and PN, which we introduced in Section 5.1.1.2. Apart from the performance, we carefully study the security and privacy guarantees of our designs to ensure that they indeed enable privacy-preserving company benchmarks. Thus, in Section 5.1.2.6, we discuss these aspects in detail. Finally, we conclude our evaluation with a performance and suitability comparison of both designs in Section 5.1.2.7.

#### 5.1.2.1 Design Overview: Company Privacy and Algorithm Confidentiality

The main difference between our designs lies in the underlying private computing concept (hardware- vs. software-based), i.e., we either utilize TEEs (cf. Section 2.3.2) or build on HE (cf. Section 2.3.1). While TEEs can retain inputs and computed KPIs

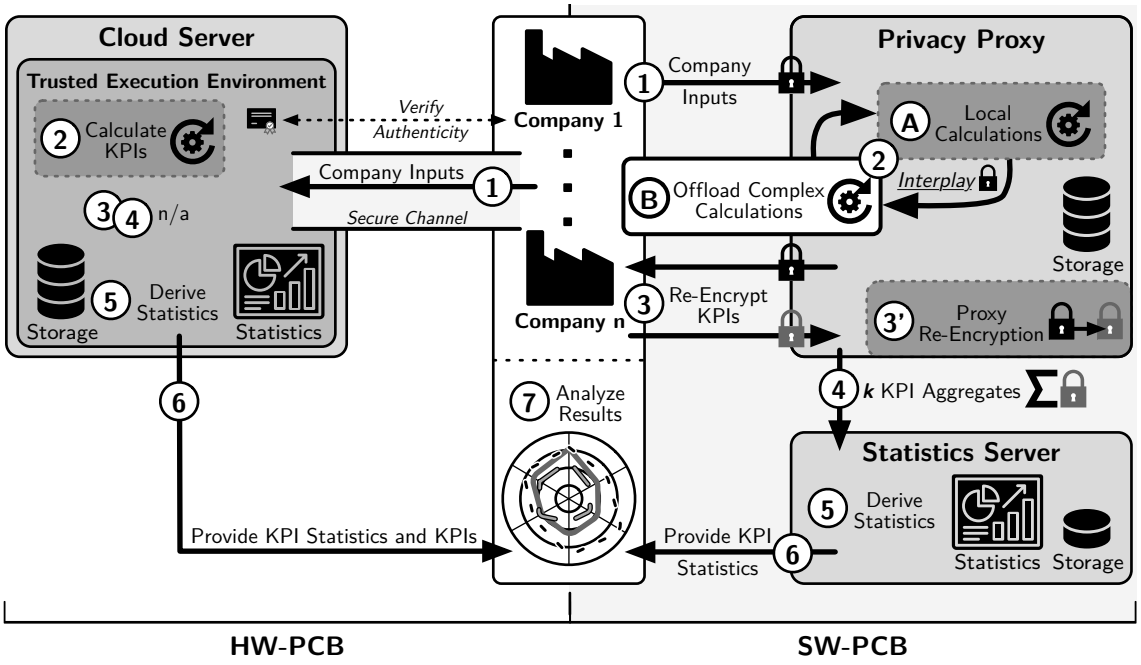



**Figure 5.3** Our hardware security-based approach HW-PCB (left) can be realized with a single server. However, it requires a TEE. Contrary, our software security-based approach SW-PCB (right) uses two non-colluding servers to ensure the confidentiality of inputs and the valuable algorithm. Eventually, companies analyze and optionally address the results of the benchmark.

of each company within the protected enclave in HW-PCB, SW-PCB's privacy proxy only operates on homomorphically-encrypted data, and the statistics server only has access to aggregates. Designing and evolving the actual benchmarking algorithms is entirely independent of our designs, which focus on securing the *operation* of benchmarking algorithms. Thus, the development of benchmarking algorithms remains unchanged. Conceptually, the logical steps to compute a benchmark are identical in our designs, and the overall steps are largely comparable. However, the individual realizations differ significantly. Thus, in this overview, we provide a high-level description and present all the design-specific details in Sections 5.1.2.2 and 5.1.2.3, respectively. We visualize both designs on a conceptual level in Figure 5.3.

In ①, the participating companies share their inputs with the benchmarking service. They can trigger this step at their own discretion and do not have to remain available while other companies participate in the benchmark (as expected in **G-B5**). In HW-PCB, the companies send their sensitive data through a secure channel directly into the TEE. In contrast, SW-PCB requires the participants to homomorphically encrypt their inputs with their own public keys. Subsequently, in ②, using the analyst's algorithms (**G-B2**), the KPIs are computed. While HW-PCB operates directly on plaintext data within the TEE, SW-PCB deals with HE ciphertexts: Thus, in SW-PCB, depending on the operation, the computation is either Ⓐ performed (locally) on the privacy proxy if supported by the HE scheme or Ⓑ it is offloaded to the participant. We refer to offloading as the process where the participant receives the operation and ciphertext(s) from the privacy proxy to (i) decrypt the input ciphertext(s), (ii) compute the operation on the decrypted plaintexts, (iii) homomor-

phically encrypt the result, (iv) and return it to the privacy proxy. Thereby, we circumvent the restricted set of HE-supported computations on ciphertexts and enable analysts to include arbitrary operations in the benchmarking algorithms (**G-B5**).

Steps ③ and ④ are only relevant for SW-PCB as HW-PCB directly operates on plaintext data and KPIs within the protected enclave, i.e., no additional security measures are needed. First, to enable their aggregation, the KPI ciphertexts have to be re-encrypted with the statistics server's key. Depending on the underlying HE scheme, we can either ③' utilize proxy re-encryption directly on the privacy proxy or ③ we have to offload the re-encryption to the company. In the predecessor of SW-PCB, we only proposed the latter option [PSF+20], i.e., more operations had to be offloaded to the participating companies. Second (④), the privacy proxy aggregates the KPIs of $k$ participants that are all encrypted with the statistics server's key and forwards these aggregates to the statistics server, which can decrypt them.

The remaining steps, ⑤–⑦, are again conceptually identical for both designs. The benchmarking service derives the KPI statistics (⑤) and shares them with the companies (⑥). Finally, in ⑦, companies analyze their results to derive management decisions that they can act upon, e.g., to close possible gaps to a recognized leader.

Next, we look at the designs' specifics and our prototypical implementations.

#### 5.1.2.2 HW-PCB: Shielding the Computations

HW-PCB, our hardware-based design, utilizes TEEs to process the companies' sensitive inputs and KPIs while preserving confidentiality. This design builds on the isolation property of TEEs together with memory encryption and storage sealing to restrict the access to sensitive information to software within the enclave.

*Setup.* Since the enclave has access to company inputs as plaintext data, the setup first needs to establish trust between the running enclave, the analyst, and participating companies. This trust includes (a) the correct and benign functionality of code running inside the enclave and (b) that the enclave actually runs the intended software on a trustworthy platform. We resolve (a) by open-sourcing the enclave code, such that any interested entity can verify its functionality, and (b) via remote attestation by a trusted certificate authority. Upon successful attestation, the trusted certificate authority issues and signs an enclave-specific certificate. This certificate then serves as an enclave identifier and proves successful attestation to all entities who connect via a secure channel (e.g., TLS). Lastly, the analyst and the participating companies provision the enclave with their configuration, including the confidential benchmarking algorithm, and input data (①), respectively.

*KPI Computation.* Due to the use of a trusted enclave, the TEE, which is operated by the benchmarking service, e.g., in the cloud, may have access to all (sensitive) data in plaintext. Hence, HW-PCB natively supports arbitrary complex operations locally at the cloud server (②) and does not require any offloading (as in SW-PCB). The TEE's memory encryption ensures that the input and all intermediate computation results remain confidential, i.e., they are only accessible by/within the enclave itself.

*Aggregation.* Due to HW-PCB's computations on plaintexts, it does not require any preparatory aggregation steps (③–④). Instead, HW-PCB directly calculates the KPI statistics in the enclave (⑤). Together with their individual KPIs, in ⑥, the general statistics are sent to the companies over a secure channel. Afterward, the enclave may terminate to ensure that any data and the KPIs are no longer accessible.

*Remarks.* As a hardware-backed design, HW-PCB depends on a TEE-enabled server, which various vendors and providers offer. We exemplarily deploy it using Intel SGX.

#### 5.1.2.3 SW-PCB: Realizing Oblivious Computations

Now, we focus on the specifics of SW-PCB and the implications of utilizing HE.

*Setup.* During the setup, the analyst (cf. Figure 5.1) configures the privacy proxy (by sharing the algorithm and configuring the parameter $k$). Moreover, the statistics server generates an HE key pair that is used to compute the aggregates in ④. Finally, each participant must generate an HE key pair as well (used in ①–③).

*KPI Computation.* The privacy proxy tries to compute as many operations on ciphertexts as supported locally (Ⓐ) to ensure algorithm confidentiality. The support for complex operations (i.e., beyond $+$,$-$, and $\cdot$) depends on the utilized HE scheme. Accordingly, unsupported operations need to be offloaded to the client (Ⓑ). Here, the analyst may configure obfuscation strategies, e.g., identifier randomization, dummy requests, or blinded computations (cf. [PSF+20]). This continuous interplay (②) between local and offloaded computations concludes once all KPIs are available.

*Aggregation.* The realization of ③ depends on the support of proxy re-encryption in the utilized HE scheme: Either ③ the KPI re-encryption (to encrypt with the statistics server's key) is offloaded to the participant (who simultaneously learns its own computed KPIs), or ③' the re-encryption is performed locally at the proxy (while the encrypted KPIs are shared to the participant for decryption). Once the KPIs of $k$ companies have been aggregated (④), these aggregates are then sent to the statistics server, which decrypts them and combines them with existing KPI statistics in ⑤. Eventually, in ⑥, the general statistics are retrievable for all participants.

*Remarks.* In SW-PCB, we have no requirements on the required hardware, as data is protected through a software-based (HE) approach. However, this design comes with limitations of the locally-supported HE operations. Furthermore, separating privacy proxy and statistics server is crucial to prevent the decryption of (unaggregated) ciphertexts that contain sensitive company inputs or KPIs. For a detailed description of PCB (the predecessor of SW-PCB), we refer to our previous paper [PSF+20, Section 4.3], which also features an elaborate sequence chart that illustrates the protocol steps of PCB [PSF+20, Appendix A.1]. When compared to SW-PCB, conceptually, the description of PCB only differs in the aggregation phase, where the (final) KPI re-encryption prior to the aggregation is interactive (i.e., it involves the participant). In contrast, in SW-PCB (in ③), we also allow for proxy re-encryption to prepare the KPIs for subsequent aggregation.

#### 5.1.2.4 Implementations and Building Block Evaluation

We created corresponding prototypes to assess whether our designs are suitable for real-world deployments, which we briefly introduce in the following. Subsequently, we discuss the performance of the utilized technical building blocks. We primarily focus on HE due to its (expected) computational overhead. After this synthetic analysis, we evaluate our designs using two real-world applications in Section 5.1.2.5.

##### Implementation and Experimental Setup

Our implementations all support the same input format for algorithms and company inputs. We utilize a translator script to dissect the (human-readable) algorithm into atomic operations [PSF+20] because the tree structure of the algorithms (with dependences between the formulas) prevents us from directly computing all formulas.

Our implementations of HW-PCB and SW-PCB are publicly available [SrcC23b].

**HW-PCB.** We implement HW-PCB in the programming language Rust and further utilize Scone [ATG+16] in version 5.7.0, which eases the use of TEEs by providing an attestation process and convenient deployment through Docker containers.

**SW-PCB.** To holistically evaluate the implications of different HE schemes, we implement two prototypes of SW-PCB that primarily differ in the utilized HE scheme.

*SEAL Version.* First, we re-implement our previous prototype of PCB [PSF+20] with the most recent version of Microsoft SEAL [Mic18], in particular, through SEAL-Python [Hue19] in version 4.0.0. We further extended this revised implementation to also support array computations, as required by our PN use case. For computational efficiency, we utilize SEAL's packing feature, i.e., we can encrypt a complete array in a single ciphertext, and array operations can be computed component-wise. Since Microsoft SEAL does not support bootstrapping (i.e., the levels must be configured according to the algorithm [PSF+20, Appendix A.2]), strictly speaking, the implemented HE scheme is not fully homomorphic. Hence, our SEAL-based prototype of SW-PCB essentially utilizes an SWHE scheme.

*CONCRETE Version.* Second, we explore the benefits of programmable bootstraps with our proof of concept that utilizes CONCRETE [CJL+20] (concrete-integer in version 0.1.0). Programmable bootstraps promise to enable computations of arbitrary univariate functions, such as the modulus or power with a scalar exponent, directly on the privacy proxy. Thus, conceptually, this HE scheme can significantly reduce the number of required offloaded operations, as we later detail in Table 5.2.

**Limitations.** HW-PCB has no implementational limitation. Concerning SW-PCB, on the one hand, our SEAL-based version does not support bootstrapping, and Microsoft SEAL does not (yet) implement proxy re-encryption. Hence, in our implementation, we have to resort to offloading in Step ③. These limitations are not conceptual but rather follow from Microsoft's implementation of CKKS [CKKS17]. Furthermore, we only communicate internally, i.e., we neither simulate constrained network links nor do we configure secure communication channels. On the other

hand, the utilized library in our CONCRETE-based version is still an early, experimental version with significant limitations: (i) The datatypes are limited to Booleans and integers (up to 64 bit), (ii) currently, it only offers symmetric encryption, and (iii) proxy re-encryption is not yet supported. With this proof of concept of SW-PCB, we want to highlight the conceptual benefits of programmable bootstraps.

**Experimental Setup.** For both designs, we run our implementations of all relevant entities, i.e., companies and server(s), on a single commodity computer with moderate resources (Intel i7-7700 with 16 GB RAM and a regular SSD). On this machine, HW-PCB utilizes Intel SGX to secure its computations within an TEE. All entities communicate over the loopback interface. We conduct 50 runs for each measurement, compute the mean, and calculate 99 % confidence intervals. For our evaluations, we rely on 128 bit-level security. In SEAL, we configure polynomial moduli of 16 384 (7 levels) and 8192 (4 levels) for IM and PN, respectively (cf. Section 5.1.1.2).

### Performance Implications of the Utilized Building Blocks

Before evaluating real-world benchmarks, we first discuss the general performance implications of utilizing private computing for privacy-preserving company benchmarking. Thereby, we obtain insights into the scalability of our designs (**G-B5**).

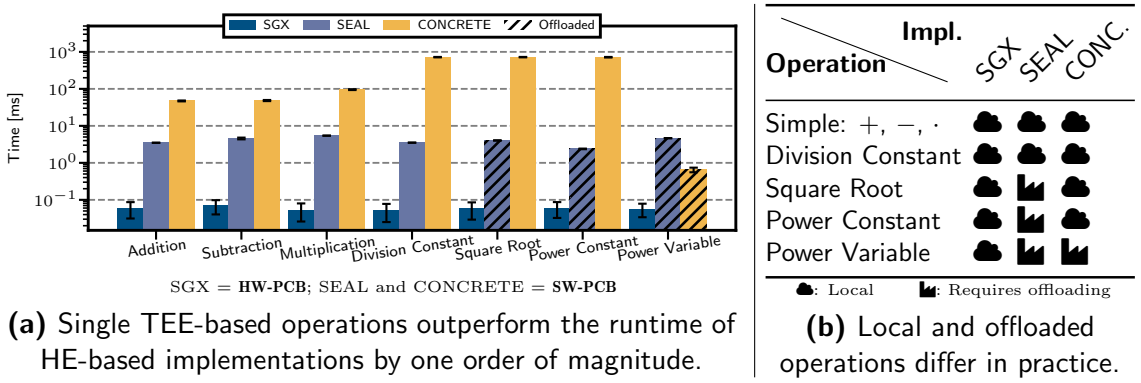
**HW-PCB.** Our HW-PCB design only differs from straightforward (insecure) computations of KPIs in the use of confidential computing, i.e., it does not introduce additional communication or computations. Hence, the overhead of utilizing a TEE in HW-PCB equals its computational slowdown compared to the insecure execution, which related work specifies with a factor of $\leq$2.5 [MPFS22]. Our measurements, covering atomic and nested computations, confirm this upper bound with overheads ranging between 25 % and 127 % [Vla22], indicating a reasonable slowdown.

**SW-PCB.** In contrast, utilizing HE to ensure confidentiality introduces performance, communication, and storage overhead. In the following, we refer to synthetic measurements from prior work [PSF+20] to summarize the full extent of this overhead.

*Atomic Operations.* First, concerning single operations [PSF+20, Figure 4], HE introduces a 5-fold increase in runtime. When offloading computations, this overhead may further decrease depending on the underlying network link. For the computation of two numbers, we measure a significant network overhead (more than 5500 fold) due to the size of the transferred HE ciphertexts in comparison to plaintexts. Moreover, offloaded computations require the transfer of additional ciphertexts between the privacy proxy and the participant, further increasing the overhead.

*Nested Computations.* Second, we observe a linear correlation between an increasing chain length and the total runtime [PSF+20, Figure 5]. In general, local computations at the proxy outperform all offloaded operations. The benefit of local computations is especially apparent for the network overhead as no (large-sized) ciphertexts need to be sent back and forth between the privacy proxy and the participant.

Naturally, the number of KPIs and the required operations to compute them amplify this overhead. However, since the KPIs (as well as the number of participants)




**(a)** Single TEE-based operations outperform the runtime of HE-based implementations by one order of magnitude.

**(b)** Local and offloaded operations differ in practice.

**Figure 5.4** Apart from confidentiality benefits, the implications of computing operations locally (in contrast to offloading them) have a performance impact in HE-based implementations.

are independent of each other, they do not add polynomial complexity. Finally, the ciphertext sizes also introduce storage overhead as the privacy proxy has to (temporarily) persist the company's inputs, intermediate results, and the final KPIs.

*Aggregation Overhead.* The separation of the server into two conceptual entities (privacy proxy and statistics server) introduces network overhead by design. This overhead is tunable, i.e., when increasing $k$, thus aggregating more results, fewer ciphertexts have to be transferred. However, in comparison to the repeated ciphertext transfer between the privacy proxy and the participants, this overhead is negligible.

**Comparing HW-PCB and SW-PCB.** To further assess the general performance, we now look at the setup and run times. We observe that the setup times are negligible ($17.893 \pm 0.015$ s for HW-PCB and $3.424 \pm 1.164$ s for SW-PCB). Looking at the run times, we compare the performance of single operations in Figure 5.4a. HW-PCB is one order of magnitude faster than our SEAL- and CONCRETE-based implementations of SW-PCB, and it does not require any offloading (Figure 5.4b). The performance of the CONCRETE-based version will likely deteriorate in the future when supporting other datatypes, such as floats or larger integers. Still, we want to emphasize the potential of programmable bootstraps, i.e., the ability to compute additional complex operations without offloading. While the overhead of computing HE ciphertexts at the privacy proxy is already significant, offloaded operations further slow down SW-PCB; especially for constrained network links.

After this discussion of the conceptual overhead in both designs, we continue with our evaluation of two real-world benchmarks in the next subsubsection.

#### 5.1.2.5 Evaluating Real-World Benchmarks

Our two real-world applications guided our evaluation to assess the suitability of our designs for real-world use, as we detail in the following. Primarily, we focus on the performance of the KPI computation as it covers the majority of relevant operations. Therefore, we do not report any numbers on the aggregation phase (③–⑤) due to its low computational footprint, as we have presented in previous work [PSF+20].

We have further shown that the impact of varying $k$ is negligible in comparison to the overall runtime needed to compute the KPIs using SW-PCB [PSF+20].

*Performance.* When measuring the overall runtime to complete the considered benchmarks, we notice that both designs are suitable for real-world deployments as the larger IM application concludes after $0.115 \pm 0.019$ s and $634.008 \pm 0.538$ s for HW-PCB and the SEAL-based SW-PCB, respectively. For our second application PN, we measure $0.080 \pm 0.001$ s and $34.409 \pm 0.044$ s. Thus, regarding scalability (**G-B5**), analysts could even offer significantly larger benchmarks while maintaining the confidentiality of sensitive information and the benchmark algorithm. For example, Kerschbaum [Ker08] outlines that benchmarks could consist of up to 200 KPIs. Overall, the runtimes for single operations (cf. Figure 5.4a) amplify in real-world benchmarks. Hence, the performance of SW-PCB remains inferior to HW-PCB.

*Accuracy.* While HW-PCB satisfies the exactness (**G-B4**) by design (the computations are simply moved in the enclave), our SEAL-based SW-PCB uses approximate arithmetic, i.e., when processing floats, we encounter precision losses. As we perform computations on approximated numbers, the precision loss amplifies, especially for long chains of operations. When using SW-PCB and suffering from insufficient accuracy, the benchmarking algorithm could be tweaked to better fit the precision of the utilized HE scheme. So far, we have studied real-world benchmarking algorithms that have not been designed with HE-induced inaccuracies (i.e., approximated representations and precision loss) in mind. In this regard, the analyst could, for example, adapt the algorithm by scaling numbers or replacing imprecise operations. Nonetheless, even without any tweaks, our SEAL-based SW-PCB is feasible (**G-B4**). Overall, we observe only minor deviations: $4.0 \pm 0.3$ % for IM and 0.00 % for PN.

*Ciphertext Overhead.* As we have discussed in Section 5.1.2.4, HW-PCB does not introduce noteworthy storage and network overhead by design. Thus, we now focus on SW-PCB. Relying on HE introduces storage and network overhead due to comparably large ciphertext sizes. With our configured polynomial modulus, the ciphertext size for IM is at most 1.842 MB, i.e., even the up- and download of thousands of ciphertexts (when sharing inputs or during offloading) is feasible. In previous work [PSF+20, Section 5.2.2], based on IM, we have already shown that the runtime is practical, even with bandwidth-constrained network links. Consequently, ciphertext overhead does not prohibit real-world applications of SW-PCB.

Moving on, we discuss our designs' security before comparing them in detail.

#### 5.1.2.6 Security Discussion

After studying the performance of our designs, we now discuss their security (guarantees) in real-world deployments. Specifically, we primarily discuss the confidentiality-related goals of company privacy (**G-B1**) and algorithm confidentiality (**G-B3**).

We consider malicious-but-cautious entities (cf. Section 2.1.2.1) in our research and further assume authenticated communication. We envision that an industry association operates—funded by membership fees—the statistics server in SW-PCB.

Naturally, our work bases on the security of the established secure communication channels, the used (technical) building blocks, and properly chosen key lengths (ensuring an adequately secure mode of operation).

*Misbehavior.* Companies most likely have to pay for their participation in company benchmarks, discouraging impulsive or destructive behavior. Submitting incorrect inputs is disincentivized as this behavior equals a loss of the participant's investment as their computed KPIs are skewed along with the general statistics (i.e., average, minimum, and maximum). If the analyst fears that companies might pay to deliberately render the insights of the benchmark useless or phony for their competitors, she could dispatch employees who observe the participants' behavior on-site.

Next, we individually discuss the security of our proposed benchmarking designs.

**HW-PCB.** The security of HW-PCB builds on hardware-based security. Consequently, the hardware vendor must be trusted, i.e., it serves as the root of trust. Using remote attestation (a key feature of TEEs), we can establish a trust chain to the enclave and the code running within the enclave. Hence, participants only have to verify this chain and the running code using certificates and cryptographic signatures. If the security has been correctly attested, all information and computations are shielded within the TEE, ensuring **G-B1** and **G-B3**. Thus, HW-PCB is secure by design. However, the multitude of (past) vulnerabilities in TEEs [FYDX21] could negatively impact the trust in this technology. Consequently, we consider SW-PCB to be a sensible alternative that does not build on hardware-based security.

**SW-PCB.** This design protects the company inputs, intermediate results, and KPIs using HE. Its security builds on the privacy proxy and statistics server not colluding. Then, the privacy proxy never has access to any (sensitive) information in plaintext as it lacks the corresponding HE decryption keys. Moreover, the statistics server only receives aggregates of $k$ participants, i.e., it cannot deduce any details about specific companies if $k$ is reasonably large (i.e., $k > 3$) [PSF+20]. Hence, sensitive company data is protected (encrypted) at all times, which satisfies **G-B1**.

In both designs, contrary to related work, the benchmarking algorithm is never shared with the participants by design (**G-B3**; the cloud server in HW-PCB and the privacy proxy in SW-PCB can be operated by the analyst to avoid any trust in third parties). However, in SW-PCB, we need to offload specific complex operations (cf. Section 5.1.2.3). Thus, fractions of the algorithm, along with their intermediate results, need to be shared with the participants, slightly violating the intended algorithm confidentiality. The benchmarking service can utilize different obfuscation strategies [PSF+20], such as dummy requests, element blinding, and request randomization, or—at the expense of reduced accuracy—an approximation of sensitive operations on the privacy proxy to mitigate the implications of offloading.

*Choice of $k$.* The privacy proxy only forwards computed KPIs to the statistics server after $k$ companies participated, preventing linkings between KPIs and companies. Thus, configuring the aggregation parameter $k$, and thus the size of the anonymity set, is a use case-specific trade-off weighing company privacy, flexibility, and the number of participants. We refer to our previous paper [PSF+20, Section 4.3.2] for an elaborate discussion of how to choose $k$ in real-world deployments.

| Criteria \ Design | HW-PCB SGX | | SW-PCB SEAL-based | | SW-PCB CONCRETE-based | |
|---|---|---|---|---|---|---|
| **Setting** | IM | PN | IM | PN | IM | PN |
| **Performance** | ★★★ | | ★★☆ | | ★★☆ | |
| ▶ Setup | Remote attestation | | Exchange of key material | | | |
| ▶ Run Time [s] | $0.11 \pm 0.02$ | 0.08 | $634.0 \pm 0.5$ | 34.4 | Unknown | |
| **Accuracy Loss** [%] | Exact | | $4.0 \pm 0.3$ | 0.0 | Unknown | |
| **Ciphertext Overhead** | ★★★ | | ★☆☆ | | ★★☆ | |
| ▶ Offloading [#] | None | | ↓1487 ↑745 | ↓53 ↑28 | ↓84 ↑42 | ↓0 ↑0 |
| ▶ Networking [≤×MB] | No overhead | | 1.842 | 1.053 | Unknown | |
| **Ease of Use** | ★★★ | | ★★½ | | Unknown | |
| **Security** | ★★★ | | ★★☆ | | ★★★ | |
| ▶ Assumptions | Trusted hardware | | Secure HE scheme | | | |
| ▶ Trust in Participants | Not required | | Non-collusion required | | | |
| ▶ Confidentiality Issues | None | | Minor (offloading) | | Barely any (offloading) | |
| ▶ Own KPIs | After/with aggregation | | Before aggregation | | After/with aggregation | |

**Table 5.2** Comparison of our hardware- and software-secured benchmarking designs.

*Entity Collusion.* In the following, we discuss several combinations of entity collusion.
–Company and Another Entity. Due to the HE ciphertexts, a company and the privacy proxy or the statistics server cannot jointly compromise any secret data in SW-PCB.
–Privacy Proxy and Statistics Server. If privacy proxy and statistics server collude, they can only decrypt the aggregated KPI ciphertexts in deployments that offload the re-encryption to the company in ③. Additionally, they can map averages to companies that are part of an anonymity set with size $k$. In implementations that utilize proxy re-encryption, they can decrypt all company inputs, intermediate results, and KPIs. Moreover, a misbehaving analyst could further define company inputs as KPIs and decrypt the "computed" KPIs if he colludes with the statistics server. This fundamental problem exists for any approach where (i) the analyst can freely define the algorithm and (ii) the participants cannot judge the importance of an input for the benchmark. HW-PCB counters this issue of algorithm confidentiality (**G-B3**) by supporting the use of attested enclaves (and running source code).
–Multiple Companies. If $k$–1 participants collude simultaneously, they can potentially reconstruct the KPIs of the non-colluding company based on the (public) KPI statistics and their own KPIs. However, this action is a punishable offense as it clashes with cartel law [SH13]. Furthermore, such an attack is unrealistic for benchmarks with many participants and can be easily mitigated with a sufficiently large $k$.

Consequently, SW-PCB ensures the privacy needs of real-world benchmarks.

#### 5.1.2.7 HW-PCB vs. SW-PCB: Selecting the Best Design for Deployment

We compare both diametrical security concepts when realizing benchmarking information systems in Table 5.2 to give a concise overview and to allow for well-founded deployment decisions. In the following, we briefly summarize the specific properties.

**Performance.** Setting up our designs is a one-time task and thus negligible with times below 18 s. Given that company benchmarks are not an everyday task, the

runtime for each company is more than suitable for real-world applications, even with significantly larger benchmarks. Our real-world applications further underline this claim (IM: $634.008 \pm 0.538$ s and PN: $34.409 \pm 0.044$ s). Even when quadrupling the number of KPIs in IM (to reach around 200 as expected by Kerschbaum [Ker08]) or when dealing with a constrained network link, our benchmarks conclude in less than a day, which is a considered boundary in related work [Ker10]. The TEE- and HE-induced overheads are reasonable in light of the confidentiality benefits.

**Accuracy.** HW-PCB features exact computations on plaintexts by design. In contrast, the loss of precision for SW-PCB is tolerable as (i) the deviations affect all companies and (ii) benchmarks primarily concern the relative positioning [MdRC12]. Moreover, the evaluated real-world algorithms have not been tailored to the use with HE. Given that the inaccuracies mainly follow from small numbers [PSF+20], the analyst could easily scale the inputs and formulas to mitigate such deviations.

**Ciphertext Overhead.** In addition to the overhead of working with ciphertexts in SW-PCB, we further have to rely on offloading to compute a subset of complex computations locally at the companies. Recent advances (e.g., CONCRETE [CJL+20]) even promise to reduce the required offloading. Irrespective of such advances, companies receive more ciphertexts than they send, which fits the imbalance of today's Internet connections. Considering ciphertext sizes of 1.842 MB for IM, uploading 1.391 GB per company is feasible, even with bandwidth-constrained network links.

**Ease of Use.** For HW-PCB, the ease of use for participants is comparable to the traditional participation in unsecured benchmarks. Accordingly, companies can easily input and query their information using a website. In contrast, SW-PCB depends on the availability of cryptographic libraries to enable the encryption of sensitive information. Given the complexity in deploying and configuring corresponding libraries [VJH21], in previous work [PSF+20], we demonstrated the feasibility of running SEAL-based implementations of SW-PCB in regular web browsers. Using WebAssembly [HRS+17], a binary code-based language that enables platform-independent execution of low-level code in web browsers, we created a web-based client for participating companies. Compared to a native SEAL-based implementation, we observed an expected [HRS+17] overhead of around 15 %. Regardless, we are confident that the improved ease of use outweighs the decreased performance.

Concerning reoccurring operational costs, HW-PCB only requires a server with TEE support, which is commercially available at major (cloud) vendors. In contrast, SW-PCB does not introduce specific hardware requirements, but its operations are computationally more expensive. In real-world deployments, analysts could operate the cloud server and privacy proxy, respectively, and fund them through participation fees. If needed, our designs support scaling out the cloud server and privacy proxy, respectively, e.g., to support a tremendous number of participants. Aside from that, industry associations could fund the statistics server using their membership fees to prevent collusion attacks [PSF+20]. Generally, HW-PCB is cheaper to operate with fewer overheads *if TEEs are trusted* by all entities, compared to HE-based SW-PCB.

**Security.** HW-PCB requires specific hardware for its operation and is secure and privacy-preserving if the trusted hardware is realized as intended. Given that com-

panies establish a secure tunnel into the enclave, HW-PCB reliably protects the algorithm, all inputs, and the computed KPIs at all times. In contrast, SW-PCB does not depend on specifically-trusted hardware but on secure and properly-configured HE schemes. We further require non-collusion among privacy proxy and statistics server to ensure company privacy. Supported obfuscation strategies can help to prevent offloading-induced information leaks (companies have access to the operator and intermediate data). As indicated before, modern HE schemes promise to further reduce the required offloading. Finally, while our SEAL-based SW-PCB uses offloading for the KPI re-encryption as part of ③, which enables companies to abort the protocol (then, they only have access to their KPIs), implementations that support proxy re-encryption (③') provide the same guarantees in this regard as HW-PCB.

**Flexibility.** By design, HW-PCB supports arbitrary operations within the enclave that directly utilizes plaintext data, i.e., analysts may rely on complex functions. In contrast, in SW-PCB, the locally-supported operations at the privacy proxy depend on the utilized HE scheme, as we have also summarized in Figure 5.4b. In addition to using a single scheme, SW-PCB could conceptually apply different HE schemes for parts of the benchmarking algorithm to capitalize on their respective benefits. Then, when designing the algorithm, the analyst must keep in mind that the ciphertexts of different HE schemes are usually incompatible. To address such incompatibilities, the analyst could utilize conversion algorithms, which emerged recently [CDKS21], to translate HE ciphertexts from one scheme to another. In this context, the analyst needs to carefully consider conversion overhead and complexity overhead of designing corresponding algorithms to the benefits of using different HE schemes in SW-PCB.

**Conclusion.** Our evaluation shows that concepts from private computing are readily available to secure company benchmarks in real-world deployments. Thus, when designing corresponding secure collaborations, the key question is which conceptual technology should serve as the root of trust, i.e., trusted hardware (a TEE) or an HE scheme, mainly because the remaining properties do not prohibit practical realizations, as we briefly summarize in the following. Both designs fulfill the needs of real-world benchmarks performance-wise, with HW-PCB computationally outperforming SW-PCB. While HW-PCB's accurate computations promise quick and precise results, SW-PCB is easier to deploy as it is designed for untrusted hardware (despite requiring two entities, i.e., privacy proxy and statistic server). The exact realization (design) then likely depends on the availability of a TEE and the willingness to build on its associated security assumptions (e.g., trusting the underlying security concept, the vendors, and remote attestation). Otherwise, HE-based implementations also promise secure and practical company benchmarks.

### 5.1.3 Takeaways and Future Research

In this section, we have detailed and evaluated two conceptual approaches to realize privacy-preserving company benchmarks in industry. With this contribution, we enable benchmarking participants to reliably analyze their position in the field. Using the benchmark's results, they can identify shortcomings (in relation to competitors

or recognized leaders) and derive actions to improve their (future) performance, such as adapting their current processes. We conclude this contribution by first discussing the suitability of our selected building blocks in Section 5.1.3.1. Subsequently, in Section 5.1.3.2, we briefly comment on the (expected) impact of our proposed designs. Finally, in Section 5.1.3.3, we outline potential next steps and future work.

#### 5.1.3.1 Suitability of Selected Technical Building Blocks

Benchmarks that operate with closed data and a closed algorithm require technical guarantees to properly secure this type of industrial collaboration. As we have outlined in Section 5.1.1.4, other commonly-applied approaches like secure multi-party computation are either not flexible enough to allow for (time-)independent participation or fail to sufficiently protect the valuable benchmarking algorithm. Therefore, we study two fundamentally-different concepts from private computing. First, in HW-PCB, we utilize TEEs from confidential computing to provide technical confidentiality guarantees during the benchmark to participants and the analyst alike. Second, in SW-PCB, we protect the company inputs and computed KPIs using HE and prevent the linking of KPIs to specific participants by applying the concept of $k$-anonymity. At the same time, the confidential algorithm is never shared with third parties or participants. Our evaluation and comparison of these designs underline that privacy-preserving company benchmarks are realizable using TEEs and HE. Thus, for this practical application, we provide a profound overview of the suitability of different concepts from private computing. In the long run, we look forward to practical advances and evolutions that follow from matured programmable bootstraps, which promise an improved implementation of our SW-PCB design.

#### 5.1.3.2 Nudging Stakeholders Toward Company Benchmarks

With this contribution, we present readily-available and real-world-applicable designs for company benchmarking that account for the confidentiality needs of involved stakeholders. Just like related work that focuses on privacy-preserving company benchmarks, we account for the sensitivity of company inputs and computed KPIs with technical guarantees. This way, we address prevalent privacy concerns of cautious stakeholders, encourage them to participate, and eventually improve the usefulness of the benchmark for all participants. Moreover, in contrast to related work, we also consider the analyst's confidentiality needs. Thereby, we improve the overall situation of company benchmarks in industry in two ways. First, with sufficient confidentiality in place, we nudge analysts to invest additional effort in deriving meaningful KPIs, which promises to also improve the usefulness and impact of the benchmark. Second, we most likely encourage additional analysts to offer company benchmarks as our designs reliably protect their competitive advantage and intellectual property. Consequently, our contribution fosters comparisons across supply chains that allow companies to identify unrealized potentials in their operations, effectively supporting the evolution of the industrial landscape.

#### 5.1.3.3 Future Work and Next Steps

As we have stressed (cf. Section 5.1.2.7), our designs are suitable for real-world deployments: They are sufficiently performant (i.e., easily conclude within one day per company) while ensuring the confidentiality of sensitive information. Thus, most prominently, we look forward to real-world use of our work. Given the profound level of readiness of our designs, we see no immediate need for future work. When having the potential offloading limitations of SW-PCB in mind, future work could study approaches that mitigate any remaining information leaks along the following aspects: (i) evaluate approximation functions that approximate complex operations that are unsupported by the utilized HE scheme, (ii) empirically investigate the benefits of obfuscation during offloading, and (iii) in the long run, follow the improvements that stem from programmable bootstraps (e.g., CONCRETE). Apart from company benchmarks, we are interested in following the rapid evolution of private computing (including its building blocks) and its implications on other secure collaborations, both in the context of privacy-preserving comparisons and beyond.

This subsection concludes the presentation of our third contribution, which concerns company benchmarking with known collaborators across supply chains. To account for the confidentiality needs of all involved stakeholders, we have proposed two designs that build on TEEs and HE, respectively. In the following, we move to our fourth and final contribution, which explicitly considers information sharing with unknown collaborators across supply chains. In such settings, collaborations are decisively challenged by the lack of established trust among the involved entities.

## 5.2 Privacy-Preserving Parameter Exchange

For our fourth contribution, we focus on the exchange of information across supply chains where collaborators are unaware of each other and their respective information. Hence, we specifically have to deal with the challenge of privacy-preservingly sharing information with competitors or strangers. Companies are generally interested in retrieving valuable insights and knowledge from the industrial landscape to improve their production (processes). However, confidentiality concerns challenge this desire because stakeholders are unwilling to publicly disclose sensitive information. Consequently, stakeholders demand secure collaborations that ensure an accurate and privacy-preserving matching of said information. Moreover, simply applying unvalidated information to PPC or running processes can automatically entail unwanted consequences. Thus, in contrast to privacy-preserving comparisons, the real-world implications of utilizing privacy-preserving matchings are more severe.

To address this research gap, in Section 5.2.1, we first outline the challenges of matching and sharing information across supply chains. In addition to the sensitivity of shared information, we also have to account for the sensitivity of client queries, i.e., requests (as part of the matching) have to remain confidential as well. Hence, this setting exceeds the scope of private information retrieval. Afterward, in Section 5.2.2, we introduce our new designs for privacy-preserving exchange platforms in the IIoT. Finally, in Section 5.2.3, we conclude the presentation of our fourth contribution.

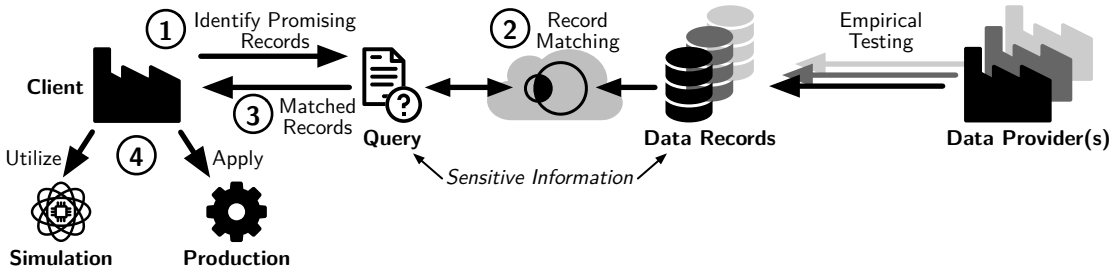



**Figure 5.5** Clients are interested in receiving sensitive business information to improve their production processes. We illustrate this setting using an application in injection molding.

## 5.2.1 Challenges for Information Sharing in Industry

To prepare for our proposed designs, in Section 5.2.1.1, we introduce the scenario of exchanging sensitive information across supply chains along with its associated challenges. Subsequently, we outline the benefits of said information sharing based on two applications in Section 5.2.1.2. Based on this overview, we then derive general design goals (Section 5.2.1.3) that capture the privacy challenges of corresponding exchange platforms. Then, in Section 5.2.1.4, we discuss related work and highlight the lack of approaches that provide sufficient confidentiality guarantees for the involved stakeholders. Finally, in Section 5.2.1.5, we present Bloom filters in more detail. This probabilistic data structure is crucial for one of our proposed designs.

### 5.2.1.1 Exchanging Sensitive Business Information

We introduce the setting of exchanging sensitive information in the IIoT based on an example that covers injection molding. Thereby, we outline how the utilization of production data and process parameters across stakeholders is desirable to evolve the industrial landscape in the future. In Figure 5.1, we illustrate a corresponding workflow that involves a retrieving company (client) and multiple data-providing companies. First, ① clients query parameters of similar processes from data providers. Afterward, ② data providers curate matching parameters from their own production and ③ send these results back to the client, which can enhance both ④ their modeling of production processes, e.g., integrating more real-world process data, as well as ⑤ their current production, e.g., utilizing well-fitting configurations.

**Need for an Exchange Platform.** While companies already gather much process information today [Kus17], which offers value for other companies as well, they lack suitable mechanisms to privacy-preservingly share this information. In particular, they are afraid of leaking sensitive business information, e.g., related to their competitive advantage and know-how. Hence, due to the manifestation of local information silos [Sch16, GPL+20] (cf. Section 1.1), the industrial landscape misses the opportunity to benefit from the matching and exchange of valuable information. Contrary to work in the medical domain [KL15, ZDJ+15, SLH+17, ZPH+17], where usually a single stakeholder offloads data to an untrusted cloud (with $m$ stakeholders querying information), we need to consider a setting where multiple stakeholders offload their data and multiple (other) stakeholders query said information. To retain the

information's usefulness, as opposed to best practices when handling sensitive user data [SCL+18], we cannot anonymize data items before or during the exchange.

Consequently, companies look for approaches that ease the exchange of sensitive business information without leaking confidential data. Simultaneously, corresponding approaches should introduce a quid pro quo for data providers to incentivize them to share their data [ZC20], even in settings with competitors. For example, companies could tap into new business models by selling their data, which they gather anyway, to third parties which themselves want to reduce their costs by utilizing this information, e.g., by optimizing processes based on exchanged information.

**Confidentiality Needs.** With business taking place in competitive environments, companies are notoriously cautious when sharing information. If at all, they are only open to sharing specifically-requested datasets [GPL+20] while retaining long-term confidentiality guarantees [Cat10], i.e., data requests must be specified precisely. Consequently, a global catalog of existing data or a way to browse available information items must not exist. Likewise, requesting companies, i.e., clients, want to utilize external information for their individual benefits, e.g., to improve their production processes, and do not want to openly share their interests. To still achieve a competitive advantage, the utilization of requested and retrieved data must remain private, including the sensitive process of identifying relevant items.

**Operational Considerations.** Given that privacy-preserving designs and security building blocks usually introduce a computational overhead and possibly add communication [WR10], the data exchange must be executed within reasonable use case-induced boundaries. This aspect is not limited to the exchange but also includes a potential setup. In particular, various security and privacy guarantees might directly contradict the feasibility of a proposed concept. Furthermore, data-providing companies might be unavailable to participate in data exchanges and their associated protocols. Hence, flexibility is needed to also account for such situations.

#### 5.2.1.2 The Usefulness of Exchanging Production Parameters in Industry

In the following, we express the benefits of exchanging information across supply chains based on two real-world applications. First, we elaborate on our previous example, i.e., an application in injection molding. Afterward, we focus on an application in the domain of machine tools. Naturally, many more use cases and domains would benefit from the availability of privacy-preserving exchange platforms.

##### Transfer Learning in Injection Molding

Given that injection molding is responsible for around 55 Mio. tons of polymer materials worldwide each year, which grosses to 16.42 % of yearly polymer production [Cer16, Pla19], improving the efficiency of corresponding production processes promises a significant large-scale impact. We refer to Sections 3.2.3 and 3.2.4 for a detailed introduction to production processes in the context of injection molding.

**Optimal Process Parameters.** In highly-complex production environments, determining an optimized set of machine parameters for the process setup is a major challenge. Suboptimal processes yield a higher scrap rate, result in lower part quality, or consume more energy during (mass) production. While arbitrary optimization by expert knowledge [BHSS12, MVS13], i.e., by trial-and-error, or process-oriented optimization by simulation [BMBJ19, HHMB19] is widespread, objective optimization can be achieved with unbiased mathematical approaches, such as artificial neural networks [LH21]. However, even these approaches usually require a broad foundation of training data, which is rarely obtainable during production, complicating the practical application of mathematical approaches. The application of transfer learning [WKW16], which transfers the knowledge from a source assignment to a target assignment, could address this issue. Recent research has shown the suitability of transfer learning in the context of injection molding [LHZ22]. Additionally, research also intensively looks into the transfer between simulation and experimental data [HH18, MTT+18, TGH+18, HJM+19]. The transfer among processes of different molded parts indicates promising results [HBKL20], suggesting a close correlation of the transfer learning success with the similarity of the assignments. Overall, the successful transfer of knowledge among processes with varying influencing parameters renders it a promising alternative to expert knowledge.

**Information Sharing in the IIoT.** Even when applying approaches like transfer learning that reduce the extent of input data, a sufficiently-large information source is needed. Unfortunately, suitable (input) data might only be available at other companies, which is currently not being utilized. Hence, a privacy-preserving exchange platform would be highly beneficial for process optimizations because its availability would enable companies to tap into several local information silos and thus build on global knowledge and experience. However, companies impose very strict usage rules [ZC20], e.g., molds are regularly owned by customers of injection molding manufacturers and only loaned to the companies for production. Therefore, any unintentional disclosure of intellectual property has to be strictly avoided.

### A Parameter Exchange for Machine Tools

When looking at the challenge of obtaining optimal process parameters in the domain of machine tools, companies usually rely on time-consuming and manual approaches (cf. Section 3.2.4). While recent developments of model-based optimization methods promise to improve this situation, as for our previous application in injection molding, corresponding optimization methods require detailed modeling of the machining process and the machine tool, which is difficult to create and not always feasible. Meanwhile, other companies may already have suitable and optimized parameters for the process at hand. Thus, the availability of a privacy-preserving exchange platform could reduce redundant work and facilitate optimizations in settings where model-based optimizations are not feasible to apply or directly applicable.

#### 5.2.1.3 Design Goals for a Privacy-Preserving Exchange of Information

Having the scenario and the usefulness of exchanging information in mind, we now derive a set of five distinct, general design goals that any privacy-preserving exchange platform in the IIoT must satisfy. These goals summarize the confidentiality needs of the participating stakeholders (**G-E1** and **G-E2**) as well as universal conceptual requirements with impact on any platform's design (**G-E3**, **G-E4**, and **G-E5**).

**G-E1: Provider Privacy.** Companies willingly offering their sensitive business information to other stakeholders have a strong desire to maintain their privacy and data confidentiality as the combined information could reveal internal information. For example, in our injection molding application, knowledge about the data provider correlated with shared geometry parameters could result in the identification of specific parts and, thereby, reveal highly-sensitive information about the implemented production processes and the company's customers. Thus, data-providing companies mandate that access to their data is only granted in parts and only to authorized parties. Furthermore, as long as data providers do not share provider-identifying information voluntarily, they must remain anonymous for all clients.

**G-E2: Client Privacy.** Protecting client requests is just as important for the success of an exchange platform. First, data providers must not be able to attribute the requested data items to the client. Otherwise, information on new developments might be identifiable and directly linked to a company. Second, the request generation, i.e., the metric used to identify meaningful data items, must remain private. In the IIoT with ubiquitous data exchanges, such knowledge constitutes the competitive advantage as the individual data items are globally retrievable.

**G-E3: Deployability.** When realizing a practical platform, two aspects are crucial. First, requests must allow for flexible matching, i.e., clients can use any metric they like to identify meaningful data items and must be able to request these identified data items. Hence, this metric can neither be part of the exchange nor should it be public during the exchange (cf. **G-E2**). Second, to incentivize data providers to offer their valuable data, a billing mechanism is required to enable new business models. Finally, data providers must not be required to remain available all the time, i.e., client requests need to be processable without their direct and active involvement.

**G-E4: Performance.** Since privacy-preserving designs usually incur performance overheads, they should scale to the needs of the respective application, and the overall runtime must be appropriate, i.e., overheads should not outweigh the introduced benefits. In this context, specifying concrete constraints is counterproductive since respective limits depend on the value of the exchanged information, i.e., very valuable data can justify significant resource needs. Likewise, introduced hardware and network requirements need to be reasonable to ensure the platform's practicality.

**G-E5: Adaptability.** Along with the previous confidentiality (**G-E1** and **G-E2**) and performance (**G-E4**) goals, the trade-off between security and performance is of crucial interest. Given that some data is more sensitive than other, it should be treated accordingly: Ideally, a concept is in place to deal with such needs.

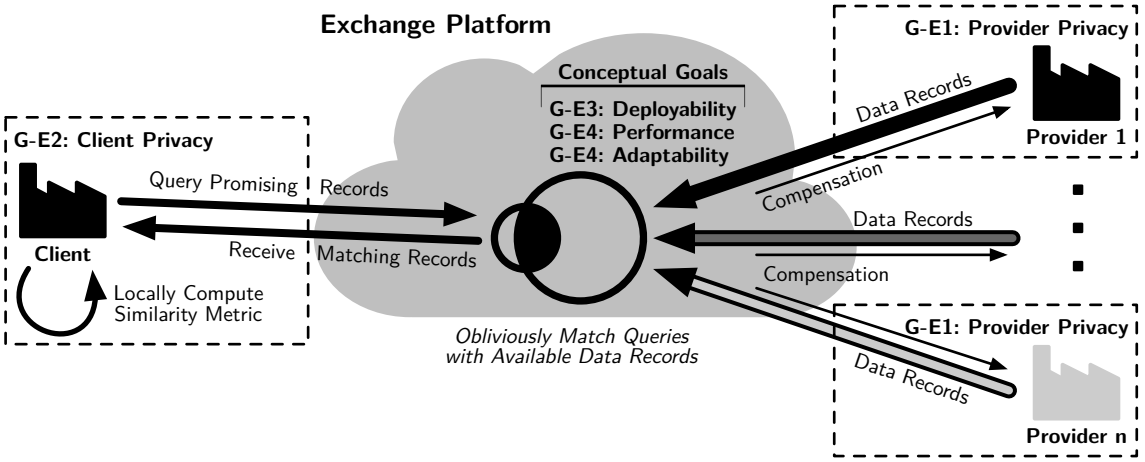



**Figure 5.6** Exchange platforms in the IIoT have to address the confidentiality needs of clients and data providers alike. Otherwise, their compliance with conceptual design goals is irrelevant.

| Proposed Approach | Client Privacy | Server Privacy | Feasibility | Trust Assumptions |
|---|---|---|---|---|
| Private Information Retrieval [CGKS95] | ● | ○ | ◕ | ● |
| Remote Knowledge System [DDM+19] | ● | ● | ◑ | ● |
| Symmetric Searchable Encryption [SWP00] | ● | ○ | ● | ○ |
| Public-Key Searchable Encryption [BDCOP04] | ● | ○ | ◕ | ◔ |
| Sharing of Sensitive Information [DCLT10] | ◕ | ◑ | ◕ | ◑ |
| Private Database Queries [BGH+13] | ● | ◑ | ◑ | ◑ |

**Table 5.3** We rate related work from ○ over ◔, ◑, and ◕ to ● to describe to which extent it satisfies the respective property and observe that no approach satisfies all aspects. The property "trust assumptions" specifically states whether the approach relies on assumptions that do not match our outlined scenario.

To establish a better understanding of these critical design goals, we visualize them in Figure 5.6. We illustrate the abstract design of an exchange platform through a cloud but stress that different realizations are conceivable (cf. Section 2.1.3).

#### 5.2.1.4 Related Work

With these design goals in mind, we now present related work on privacy-preserving information retrieval. We summarize their key properties in Table 5.3 and further discuss to which extent related work is applicable to our outlined scenario.

*Private information retrieval* [CGKS95] protocols deal with privacy-preserving data retrieval from a database. However, these protocols only consider the client's privacy, i.e., the query is hidden from the database server, while the server's privacy (**G-E1**) is not protected. Accordingly, this class of protocols [CGKS95,MH20] is not applicable to our scenario, as the client is not allowed to learn anything beyond the matching item. OTs (cf. Section 2.3.1) also represent a form of symmetric private information retrieval. While they provide a profound level of privacy, OTs alone are infeasible for transmitting large data volumes, which are common in the IIoT.

Other primitives for secure computations, such as secure multi-party computation [Lin05] and HE (cf. Section 2.3.1), are available to realize privacy-preserving information retrieval as well [ZPH+17]. However, secure multi-party computations introduce high overhead (**G-E4**) and do not reach the efficiency of purpose-driven protocols for private information retrieval [DCLT10]. Similarly, HE approaches that mimic such protocols suffer from the same inefficiency [NLV11]. In addition, supporting arbitrary metrics for the matching (**G-E3**) is infeasible with HE as it either offers only a restricted set of operations or becomes overly complex [AAUC18].

The *privacy-preserving remote knowledge system* [DDM+19] tackles the feasibility of data retrievals via OT. A PSI initially determines matching elements, such that only matched elements induce an expensive OT. However, size limitations of the PSI restrict the number of data items that are retrievable with this approach.

Both *symmetric searchable encryption* [SWP00] and *public-key searchable encryption* [BDCOP04] allow the delegation of a search operation to an untrusted third party, e.g., a cloud. These approaches encrypt data and search queries. The third party returns matched elements to the client without learning the plaintexts. Applied to our scenario, data providers could upload their (encrypted) data, and clients could then send a search query. However, both approaches assume that the party delegating the search, i.e., the client, is allowed to freely access *all* stored data without restrictions. Accordingly, they cannot satisfy the required server privacy (**G-E1**).

*Privacy-preserving sharing of sensitive information* [DCLT10] considers comparable design goals, i.e., demanding both client and server privacy (**G-E1** and **G-E2**). This approach introduces a semi-trusted third party, called isolated box, which cannot access plaintext on its own. The design requires non-collusion among client and server. All data items feature multiple attributes that allow clients to submit disjunctive queries over multiple attributes (conjunctive queries are not supported). Unfortunately, disjunctive queries are not useful in our scenario as all input parameters have to match the client's query. Additionally, the design only considers a single data source (the server), while we have to support multiple data providers. The encryption process, which is offloaded to the isolated box, requires knowledge of how many records with a certain attribute-value pair exist. Accordingly, the encryption cannot independently be outsourced from the server to the data providers (**G-E3**). Therefore, adapting this approach to our discussed scenario is far from trivial.

The approach of *private database queries using SWHE* [BGH+13] extends the previous solution with conjunctive queries. However, it also assumes that all data is provided by a single central server. Moreover, the design requires the computation of an inverted index by the server entity. Due to the fact that this computation bases on plaintext access to the stored data, an additional semi-trusted storage server cannot perform it, i.e., computing the inverted index requires information on how many records with a certain attribute-value pair exist. Hence, the challenge of adapting it to multiple independent data providers, as required by our scenario, remains. Additionally, we expect that this work does not scale to our scenario (**G-E4**) as it was only evaluated with up to 5 attribute-value pairs. In contrast, our scenario calls for more diverse query parameters, each adding a random linear combination.

**Research Gap.** This overview shows that related work has proposed many approaches for private information retrieval. However, they are inapplicable to our scenario as they either fail to provide sufficient server privacy, which is essential for secure collaborations, or cannot be adapted to our scenario without significant effort.

#### 5.2.1.5 Preliminaries: A Space-Efficient Probabilistic Data Structure

In addition to the building blocks that we have already presented in Section 2.3 (primarily OTs and PSI), we also utilize Bloom filters, a probabilistic and space-efficient data structure, in one of our designs. Therefore, we introduce corresponding background information on this established building block in the following.

**Bloom Filter.** Bloom filters allow for efficient membership tests without an efficient possibility to extract a list of all inserted elements [Blo70]. Apart from insertions, Bloom filters support membership tests that check whether a specific element was inserted. Due to its probabilistic property, such queries can return false positives with a tunable false positive rate $\varepsilon$. However, false negatives cannot occur.

A Bloom filter $B$ consists of an array with fixed length $n$ and uses $k$ hash functions $(h_1, \ldots, h_k)$ to map elements to the fields of the array. Inserting an element $x$ works by setting $B[h_i(x)] = 1 \ \forall i \in \{1, \ldots, k\}$. Querying an element $y$ equals a bitwise comparison of $h_i(y) \ \forall i \in \{1, \ldots, k\}$ with $B$. Taking the false-positive rate into account, $y$ was inserted in $B$ if all set values in $h_i(y) \ \forall i \in \{1, \ldots, k\}$ are set in $B$ as well. The false-positive rate $\varepsilon$ can be computed based on the number of stored elements $m$, the length $n$, and the number of hash functions $k$ [SSK14]: $\varepsilon = (1 - (1 - \frac{1}{n})^{km})^k$. Adjusting the Bloom filter's parameters (e.g., to reduce $\varepsilon$) influences the storage size as well as the processing time of insertions and queries.

### 5.2.2 Privacy-Preserving Exchange Platforms for Industry

In the following, we introduce multiple designs that enable the exchange of information across supply chains. In contrast to related work, our designs account for both client and server privacy to address the confidentiality needs of secure collaborations in the IIoT. Conceptually, our designs are closely related but differ in the utilized building blocks. In particular, we utilize Bloom filters, OTs, and PSIs. Consequently, the privacy guarantees and computational complexity of our designs vary slightly. Overall, we provide practical exchange platforms that enable the oblivious exchange of arbitrary information through a deterministic indexing scheme.

The presentation of our fourth contribution is structured as follows. First, in Section 5.2.2.1, we give an overview of our designs before discussing the respective protocols and operators of our exchange platform. Afterward, in Section 5.2.2.2, we introduce our prototypical implementations and elaborate on the performance implications of our utilized building blocks. Subsequently, we evaluate our designs in more detail. To this end, we first conduct a performance evaluation that also covers our two real-world applications (cf. Section 5.2.1.2) in Section 5.2.2.3. Then, in Section 5.2.2.4, we discuss our designs' security and confidentiality guarantees. Finally, in Section 5.2.2.5, we compare our designs based on the findings of our evaluation.

#### 5.2.2.1 Design Space and Designs for Sharing Sensitive Information

Overall, we propose a modular concept that results in various designs depending on the building blocks used for the matching and retrieval of data items. We name our designs according to their primary building block. Altogether, we refer to them as **B**loom filter-based **P**arameter **E**xchange (BPE), **P**SI-based **P**arameter **E**xchange (PPE), and an entirely **O**T-powered **P**arameter **E**xchange (OPE). Moving on, we first introduce a general notation to provide a more formal understanding of our work. Then, we discuss the general concept before elaborating on the involved entities and our index-based record matching that enables so-called similarity metrics. Finally, we discuss different variations in the design space of our modular concept.

##### Notation for the Exchange of Information

Our designs build on records that are indexed deterministically to enable data providers to offload their records. A record $p = x \parallel y = x_1, \ldots, x_n, y_1, \ldots y_m$ consists of a payload $y$ and a number of (identifying) parameters $x_i$ for the indexing. Here, $x$ could correspond to a part that should be manufactured at a specific machine, while $y$, for example, represents used machine settings. The respective indexing is defined by $X \rightarrow H : h_k(x'_1, \ldots, x'_n) = id_{x'}$ with a use case-specific rounding function $r(x_i) = x'_i$ to derive its input, i.e., we apply a binning to match related records to the same index. Both $h$ and $r$ are globally defined by the exchange platform. We further derive $idk_{x'} \in K$ as truncation of $id_{x'} \in H$, for the indexing of AES encryption keys $k_{x'}$, i.e., the encryption key can be derived using the identifying parameters $x_i$ only. Records can share an encryption key if $K \subsetneq H$, i.e., fewer indices are available at the key server, which also handles the mapping $(idk_{x'}, k_{x'})$. To reduce the computational overhead, a smaller set size $K$ is desirable. An encrypted parameter record $c_{x'}$ is further defined as $c = E_{k_{x'}}(p)$. The storage server maintains the respective pairs $(id_{x'}, c_{x'})$. A single index $id_{x'}$ can refer to the ciphertexts of multiple records due to the rounding with $r(x)$ (putting records into bins). By design, these ciphertexts also share their AES encryption key.

Our work introduces and supports a similarity metric $s(q)$ to identify relevant records (the candidate set $S$) based on an initial record $q$. We consider this metric to be sensitive (**G-E2**). Thus, in our designs, its use is limited to the querying companies (clients). To compute $S$, a client does not require any payload data as records are indexed with their identifying parameters $x_i$ only. Eventually, the client retrieves all records $q'$ with $id_{q'} \in S$ that are available (indexed) at the exchange platform.

Using this notation, we provide a high-level design overview of our modular concept in the following before focusing on the involved entities in more detail.

##### Design Overview

As we illustrate in Figure 5.7, we realize our designs for privacy-preserving exchange platforms by splitting ciphertexts and key material over two independent operators. To achieve the desired privacy guarantees, carefully-selected operators who may not

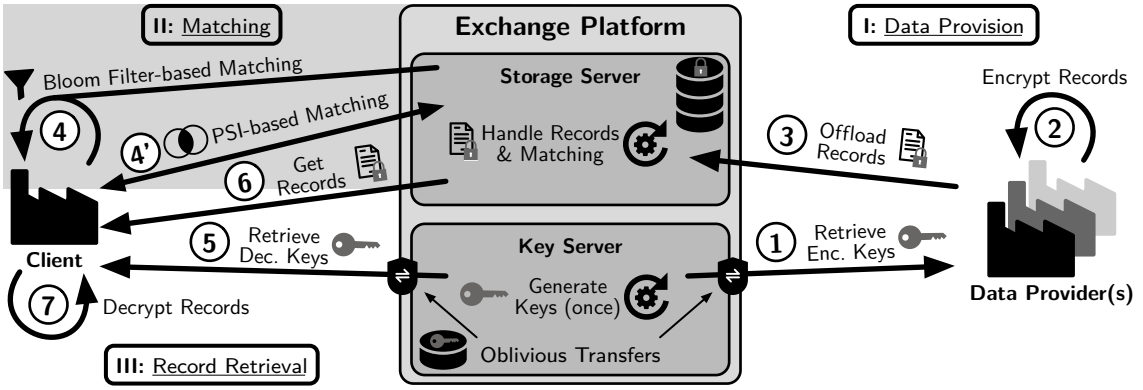



**Figure 5.7** Our designs, which facilitate the privacy-preserving exchange of information in the IIoT, are split into two components to separate the key material from shared ciphertexts.

collude must run the two servers of our exchange platform. Apart from this aspect, both clients and data providers do not have to trust any other entity. We separate the functionality of our designs into three primary phases, which we present next.

**Phase I: Data Provision.** First, ① data providers retrieve encryption keys $k_{x'}$ from the key server via OTs, ② encrypt the information they are willing to share (their records $p$), and ③ offload (cf. **G-E3**) their encrypted records $c$, annotated with the indices $id_{x'}$ to the storage server, which maintains the indices of received records (from all providers). OTs hide the providers' access patterns from the key server.

**Phase II: Matching.** The client triggers the matching phase. Then, starting with a known record $q$, the client locally computes all indices that she is interested in (her candidate set $S$) based on a similarity metric $s$ (her intellectual property). Depending on the design, the client checks the availability of these indices $id_{q'}$ either ④ through a Bloom filter or ④' using PSI. The local matching ensures client privacy (**G-E2**) because the query content ($S$) is not shared with another entity.

**Phase III: Record Retrieval.** If any records matched the client's query in Phase II, ⑤ she retrieves the corresponding decryption keys $k_{x'}$ from the key server via OT. ⑥ She further purchases the respective ciphertexts from the storage server, which also triggers the billing (out of scope for this dissertation) for this retrieval. Finally, ⑦ she decrypts the retrieved ciphertexts $c_{x'}$ and gets access to the queried data records. Again, OTs reliably hide the (client's) access patterns from the key server.

After these three phases, the client is oblivious of data-sharing providers (**G-E1**), and assuming a proper billing mechanism, the selling provider cannot identify the buyer either (**G-E2**). Furthermore, the client's valuable similarity metric is kept private as the client locally computes the candidate set (**G-E2**). Since all items are encrypted, the storage server is unaware of the mediated records (**G-E1** and **G-E2**). Moreover, the key server is oblivious of requested and transferred keys because the respective communication places via OTs (**G-E1** and **G-E2**). Finally, computationally-expensive tasks are largely performed by clients or data providers, keeping the total utilization of our platform components comparatively low (**G-E4**).

### Involved Entities

To look at our general concept in more detail, we now present the involved entities. We particularly convey their responsibilities, interactions, and trust relationships. Furthermore, we outline how our platform incorporates their individual interests.

**Data Provider(s).** Given that potential data-sharing providers invest resources when gathering valuable records [AIGARB13, GZZL18, HJM+19] and possibly share their know-how with business partners or competitors, they are only willing to contribute against compensation [ZC20], e.g., payments, and despite a required participation overhead. Additionally, the data provider's identity and valuable provided data must be protected, i.e., no third party may get access to all records. To this end, in our designs, we separate key material and ciphertexts by relying on two non-colluding operators. Our platforms bill clients to reward the provider, i.e., data providers receive payments for their records if clients retrieve them. Finally, to ease the participation, our platforms allow providers to offload data once, which is not time-critical, and further supports adding additional records at a later time.

**Client(s).** The privacy interests of clients are twofold in our scenario. First, similarity metrics are potentially valuable as they originate from ongoing research [TGH+18, HJM+19] and, thus, must be protected accordingly. Second, the initial input for the metric (i.e., a known record) is sensitive as well since it might reveal internal information [CFAF17], e.g., production plans. Apart from such privacy interests, clients should only have to pay for retrieved data records to compensate providers.

Our designs ensure local matchings, i.e., the similarity metric $s$ as well as the initial input $q$ remain at the client. Moreover, since the storage server is unable to decrypt or identify the requested records, it cannot draw conclusions from the transmitted indices. Likewise, client and key server only interact via OTs for potentially leaking requests, i.e., the key server never learns anything about the client's query. Although, depending on the number of filled indices and the used similarity metric, the matching can become time-consuming, it is usually not very time-critical. For instance, injection molding productions are planned weeks in advance [DN05] and, thus, a processing of multiple days for the matching and retrieval remains practical.

**Key Server.** The interests of the key server operator are limited to a low computational and storage footprint. While generating the key material for every possible index temporarily features a high workload and forces the server to store all generated keys, the number of keys is limited by the used OT set size. Thus, the key generation neither produces significant overhead nor requires excessive storage. Although data transfers via OTs are known to be computationally expensive and time-consuming [ALSZ13], a fundamental requirement is to meet the interests of key-retrieving providers and clients, i.e., prevent any information leakage from these entities, including the number of transferred keys [CT05]. Hence, except for non-collusion with the storage server, in our designs, no trust in the key server is required.

**Storage Server.** Generally, the interests of storage and key server operators are aligned, i.e., low overheads are desired. The storage server maintains the record indexing, supports the matching phase of the client, and returns requested records to

them. The corresponding operator is very sensitive because he must not collude with any of the data providers and the key server. Therefore, he must be chosen carefully, as we highlight at the end of this subsubsection (and in Section 5.2.2.4). While the operator must observe the indices of offloaded and requested records to enable billing, this knowledge does not allow for conclusions on any of the sensitive information, e.g., the client's similarity metric $s(q)$. Only if the storage server operator colludes with the key server operator or the offloading data provider and therefore gains insight into requested data, conclusions about the client's candidate set are possible.

#### Record Indexing and Similarity Metrics

We rely on use case-specific rounding functions $r(x_i) = x_i'$ to index our records and to support similarity metrics. Such a rounding function can resemble a binning of related records, i.e., related records are rounded to the same value such that their indices are identical. The rounding depends on the absolute value of the input to achieve a granularity adaption because smaller changes in smaller values are expected to have a larger influence than the same absolute change on a larger value. For example, in injection molding, cooling times can easily be as low as 5 s, while melt temperatures for polypropylene polymer are above 200 °C. Here, the absolute difference between the values is far more than a magnitude, demanding a use case-specific rounding. In theory, any rounding could be implemented and used as well.

As an example, we now demonstrate a rounding approach for the input (1.21, 22.22, 333.33). Our defined function $r$, which we also apply in our performance evaluations (cf. Section 5.2.2.3), rounds each input parameter $x_i$ to a certain number of digits, here exemplified with 2 for all inputs, starting from the digit with the highest potency. This rounding yields the index (1.2, 22.0, 330.0) and demonstrates that the rounding of $x_1$ uses a finer granularity than $r(x_3)$ due to its smaller absolute value.

To illustrate the corresponding computation of the candidate set $S = s(q)$, we consider the target record $(1.2, 22.0, 330.0)$, a rounding to two digits, and a metric $s$ that computes the 10 % offset for each identifying parameter. This metric yields 1.1, 1.2, and 1.3 as possible values for the first parameter because the absolute offset is 0.12 such that 1.0 and 1.4 are not covered. For the second parameter, the metric computes 5 possible values and 7 for the third parameter. Hence, the candidate set consists of $3 \cdot 5 \cdot 7 = 105$ candidates in total. In real-world deployments, similarity metrics constitute the clients' competitive advantages, i.e., in practice, they are likely to be significantly more complex than our descriptive example.

#### Design Space and Design Variants

Our proposed exchange platforms build on a modular concept that supports different building blocks to achieve different levels of technical security guarantees. In particular, we can apply different concepts to realize record matching and record retrieval, respectively. In this dissertation, we focus on applying Bloom filters and PSIs to match the records of the storage server with the client's similarity set $S$. Likewise, for the record retrieval from the storage server, we distinguish between

| Retrieval \ Matching | Bloom Filter-Based | PSI-Based |
|---|---|---|
| **Hash Key-Based** | BPE | PPE |
| **OT-Based** | N.N. | OPE |

**Table 5.4** Our modular concept for exchange platforms supports different approaches for record matching and record retrieval. In this dissertation, we discuss BPE, PPE, and OPE.

hash key-based and OT-based retrievals. We summarize the different combinations of both dimensions along with our introduced labels for them in Table 5.4.

While Bloom filter-based matchings and hash key-based record retrievals promise improved performance, they introduce less strict confidentiality guarantees compared to their PSI-based and OT-based counterparts. In this dissertation, we discuss BPE and PPE because they provide a good trade-off between performance and privacy guarantees. In contrast, OPE provides strong confidentiality guarantees but significantly limits the number of supported indices and record sizes. Thereby, OPE does not scale to the dimensions of real-world applications in the IIoT.

**BPE: Our Bloom Filter-Based Approach.** In this design variant, the storage server also inserts the indexed records in a Bloom filter. In the matching phase, the client receives the Bloom filter from the storage server, hashes all indices in the similarity set $S$, and locally tests them for membership in the Bloom filter. With this approach, the storage server only shares a probabilistic data structure of all inserted hashes and not the values or full indices. After a match, the client retrieves the corresponding data records using the hashed indices of matched records.

**PPE: A PSI-Based Approach.** To prevent potential information leaks through the Bloom filter, i.e., a list indicating all available indices shared with every client, we also propose a design variant that replaces the Bloom filter-based matching with a PSI. By using PSIs, clients only learn the matching elements. Consequently, the PSIs-based record matching improves provider privacy (**G-E1**). However, due to the limited supported size of the candidate set $S$ in PPE, we generally favor BPE over PPE despite its weaker provider privacy. In settings with specific privacy needs and comparable small candidate sets, PPE can be a suitable, more secure alternative.

**OPE: Fully OT-Powered Approach.** Given that the storage server learns the identifiers of retrieved records, the hash key-based record retrieval impairs the goal of client privacy (**G-E2**). To mitigate this effect, we allow for a record retrieval that equals the key retrieval, i.e., clients retrieve all sensitive information using OTs. Just like PPE, this design variant utilizes a PSI-based matching as well. The resulting approach OPE is conceptually similar to the work by Dahlmanns et al. [DDM+19], i.e., first conducting a PSI-based matching and then retrieving the records via OTs.

However, relying on OTs for the retrieval introduces significant limitations. Modern libraries can only transmit 128 Bit per OT because OTs have mainly been designed for the transmission of key material and not the payload itself [ALSZ13]. Since our ciphertexts are significantly larger, the applicability of an OT-based record retrieval

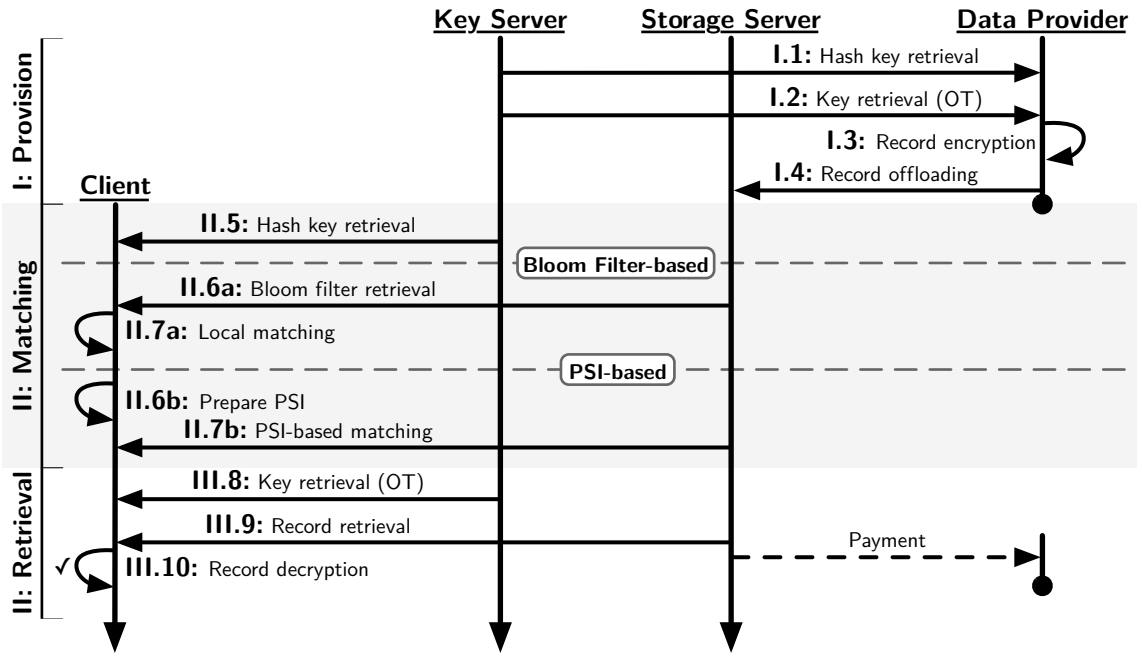



**Figure 5.8** Sequence chart detailing the conceptual steps of our exchange platform. While the matching phases of BPE (a) and PPE (b) differ, the remaining protocol is unchanged.

is limited. For example, if $t$ subsequent OTs are needed to retrieve a single ciphertext, these additional OTs increase the overhead of the retrieval by factor $t$. More importantly, the OT set size further defines the number of supported indices. Hence, only a low number of records can be handled by the exchange platform.

These limitations highlight that OPE is only applicable to small scenarios with strong privacy needs, which seems to be rarely the case for applications in the IIoT. Thus, we refrain from elaborately presenting and discussing OPE in this dissertation.

### Protocol Sequence

To elaborate on the design variants, in Figure 5.8, we illustrate the protocol sequence of our exchange platform concept. We again break it down into three separate phases.

**Phase I: Data Provision.** Initially, (`I.1`) data providers request a hash key from the key server to compute all needed indices $id_{x'}$, which prevents the storage server operator from concluding the stored data by brute-forcing the indices and increases the variability of indices. Subsequently, (`I.2`) the provider requests key material $idk_{x'}$ from the key server via OT to (`I.3`) and then encrypts the records. Finally, (`I.4`) the provider sends the encrypted records and their indices to the storage server.

While a Bloom filter-based matching can utilize the index values that have been computed with $H$, a PSI-based matching requires shorter indices ($2^{128}$ in our case and larger than the chosen OT set size of $K = 2^{20}$). Thus, we introduce a third indexing $L$ with $K \subset L \subset H$, and calculate the respective truncation for $L$ using the values inserted in the sets. The client utilizes the computed candidate set $S$,

and the server takes the indices of all stored records as their *sets* for the PSI. Although, in theory, PSIs would support the intersection of sets with a size of $2^{128}$, to achieve computational feasibility, the number of elements in the set must be reduced. Notably, in contrast to OT, the input indices in $L$ are not limited by the PSI set size, reducing the chance of clients guessing matching indices and, further, due to computational effort, preventing clients from performing PSI operations with an extensive number of elements in their candidate set (e.g., to request all records).

**Phase II: Matching.** To prepare the matching, (`II.5`) the client requests the hash key from the key server. Since BPE and PPE utilize different concepts to realize the matching, the following steps of the matching phase differ slightly. For the Bloom filter-based matching in BPE, (`II.6a`) the client receives the Bloom filter from the storage server. The previously-received hash key enables the client to (`II.7a`) derive the indices of candidates, i.e., her candidate set $S$, computed by her metric $s(q)$ based on input $q$ by locally checking whether the received Bloom filter contains the respective indices. In PPE with its PSI-based matching, the client (`II.6b`) first prepares the PSI and (`II.7b`) then performs the PSI with the storage server.

**Phase III: Record Retrieval.** After the matches have been determined, the client (`III.8`) retrieves the required decryption keys $idk_{x'}$ via OTs from the key server and (`III.9`) requests the encrypted records $E_{k_{x'}}(p)$ from the storage server using the matching indices $id_{x'}$ which subsequently triggers the payment to the data provider. Finally, (`III.10`) the client decrypts the retrieved ciphertexts to obtain the records.

This step concludes the current query of (and exchange with) the client.

### Operators of the Exchange Platform

Since non-collusion among the operators of our exchange platform is essential for the confidentiality guarantees of our designs, we describe potential operators in the following. Their costs could be covered by a participation fee paid by all participants of the exchange platform. Alternatively, the operators could also introduce per-operation fees, e.g., charge for each processed key and record retrieval.

Given that the key server is oblivious of retrievals, no trust in the operator is required. Accordingly, untrusted third parties can realistically operate it. Here, startups that charge a fee for each retrieval are potential candidates. When using a trusted third party, the key retrieval (during data provision and record retrieval) could also be implemented without OTs, sacrificing the technical confidentiality guarantees. However, as we detail in Section 5.2.2.3, the matching phase is responsible for most of the runtime. Thus, we firmly recommend secure OT-based key retrievals.

The storage server is more critical for both provider and client privacy. On the one hand, this server learns the ciphertexts of stored records and the associated data providers. On the other hand, the storage server is aware of the client's matches. Therefore, public organizations, industry associations, or the government are well-suited for hosting the storage server. They could fund the operation using donations, membership fees, or taxes. Hence, they are more appropriate to operate the storage server than a (random) untrusted, potentially-unreliable third party.

This part concluded the presentation of our modular concept and the introduction to the design variants BPE and PPE. In the following, we first introduce our implementations and then discuss the performance of their underlying building blocks.

#### 5.2.2.2 Implementations and Building Block Evaluation

To assess whether our designs are suitable for real-world deployments, we created corresponding implementations, which we briefly introduce in the following. Subsequently, we discuss the scalability and performance of the utilized technical building blocks. Their performance is essential when considering the realization of large-scale exchange platforms with large numbers of records, data providers, and clients.

##### Implementation and Experimental Setup

Now, we first specify the different components and libraries of our implementations. Afterward, we detail the experimental setup for our performance evaluations.

**Implementation.** We implemented all components, i.e., the client and data provider applications as well as the servers, in Python 3. We further rely on libOTe [Rin16a] to realize OTs and select the semi-honest 1-out-of-$n$ OT algorithm KKRT16 [KKRT16]. We interact with this library using Cython [BBC+11]. We realize both servers as Flask [Ron10] applications with Celery [Sol09] as a task queue while utilizing a Redis [San09] broker. Celery workers handle the servers' endpoints. All communication between applications and the server is protected by TLS 1.2 [DR18]. The storage server relies on SQLite [SQL00] while the key server keeps all keys in memory. In BPE, we create, process, and query our Bloom filters using Pybloomfiltermmap3 [Pra16]. For the PSI-based matching in PPE, we again utilize Cython and use libPSI [Rin16b], which runs the semi-honest PSI algorithm KKRT16 [KKRT16].

Our implementations of BPE and PPE are publicly available [SrcC20].

**Experimental Setup.** For all measurements, we utilized a single server (2x Intel XeonSilver 4116 and 196 GB RAM) and performed 10 runs each. All entities ran on the same machine and communicated over the loopback interface. We measured the data volume with tcpdump [JLM88]. We noticed an unreasonably out-of-scale overhead in the (unsupported) TLS endpoints of libOTe and libPSI, forcing us to add the expected overhead arithmetically. To this end, we evaluated the TLS handshake overhead (53.94 ms) and the maximum TLS throughput (567.16 MBit/s) on our evaluation server using Flask's TLS settings (TLS 1.2, *ECDHE-RSA-AES256-GCM-SHA384*, and the elliptic curve *secp256r1*). If not stated otherwise, we included the calculated TLS overhead based on these values (hatched in our plots).

The hash key and the encryption keys are 128 Bit long each. We parallelize the Bloom filter-based matching and the OT-based key retrieval. Even though the size of $S$ depends on the client and its metric, we expect that the server set is unlikely to exceed 100 Mio. elements, and thus, we fix the PSI set size in PPE to $2^{20}$.

### Performance Evaluation of the Utilized Building Block

In our previous paper [PBL+20, Section 7.2, Appendix B, and Appendix C.2], we measured the performance and scalability of our building blocks, namely, Bloom filters, PSIs, OTs, and similarity metrics, in great detail. Next, we briefly highlight the corresponding findings that are relevant to assess our designs' practicality.

*Bloom Filter.* Even for Bloom filters with capacities of up to 1 Bil. elements (with a false-positive rate of $10^{-20}$), we measure reasonable storage sizes for use as part of an exchange platform because one-time transmissions (to clients) of less than 20 GB are realistic nowadays [Cis20]. The Bloom filter capacity correlates linearly with the Bloom filter's storage size. Notably, the size increases linearly for an *exponentially* decreasing false-positive rate. Thus, even low false-positive rates (e.g., $10^{-20}$) yield feasible sizes. Likewise, the performance of the query time, which is relevant for the matching phase as well, is mostly unaffected by both capacity and false-positive rate and only depends on the number of performed queries. By design, querying is embarrassingly parallel as the individual queries are independent of each other. In contrast, when measuring the insert times, we notice an approximately-linear increase with both increased capacity and false-positive rate. However, the data-provision phase is not time-sensitive because it constitutes a one-time task for data-providing companies, and they are unlikely to offload their records simultaneously.

With these results in mind, we fix the capacity at 100 Mio. elements for our subsequent evaluations (as required by our applications [PBL+20]) and set the false-positive rate to $10^{-20}$, which results in a comparably small Bloom filter size (<2 GB).

*Private Set Intersection (PSI).* The set size of the PSI is the primary influence on the measured performance. It scales linearly with runtime and memory usage. Consequently, the maximal supportable PSI set size for PPE depends on the available memory at the storage server (leaving aside the serving of multiple clients at once). In the following, we select a set size of $2^{20}$, which only requires $\approx 0.6$ GB of memory per PSI. We further studied the implications of bandwidth-constrained networks and noticed a significant impact on the performance. However, even for a restricted bandwidth with 6 MBit/s, the execution time for a PSI with a set size of 1 Mio. elements stays around 45 min. We consider this time to be acceptable because the exchange of sensitive *and valuable* information may take up to several days. Given the availability of various PSI protocols, the utilized protocol can be chosen according to the use case to trade off communication and computation overhead [KKRT16].

*Oblivious Transfers (OTs).* The runtime of OTs is mainly influenced by the set size (total number of keys) and the number of OT extensions (number of retrieved keys). A large set size $K$ is desirable as more distinct encryption keys can be handled by the key server, i.e., fewer records share their encryption keys. The number of retrieved keys depends on the number of sharable records at the data provider and the number of matches at the client, which are both highly use case-specific. The runtime of the OTs further scales linearly with increasing set size and the number of performed OT extensions. As for the PSI building block, bandwidth-constrained networks or large latencies delay the runtime of OTs. However, since neither the data-provision phase

nor the record-retrieval phase are time-critical, we consider the OT runtime to be satisfactory for use in our proposed exchange platforms.

For the remaining evaluations, we fix the set size at $2^{20}$, which allows for more than 1 Mio. distinct encryption keys. With our setup and this set size, retrieving 200 keys takes less than 70 s. Again, the diversity of available OT protocols allows for balancing the trade-off between computation and communication overhead [ALSZ13].

*Similarity Metrics.* Given that the similarity metric $s(q)$ is solely computed at the client, its performance does not impact the other parts of our designs. However, the number of elements, which the similarity metric $s(q)$ returns, significantly influences the overall runtime of the matching phase. For the Bloom filter, the size of the candidate set $S$ is the primary driver of this increase, with an influence on the client only. In contrast, in PPE, the size of $S$ affects the runtime of both client and storage server. For our sample application, we introduce a fine-granular rounding to allow companies to identify potentially relevant records. Here, an increase of the granularity of one parameter ($x'_i$) yields a linear increase of candidates, while a granularity increase for all input parameters ($x'_1$ to $x'_n$) results in an exponential blow-up of the candidate set $S$. The number of input parameters $n$ is further influential because $S$ grows exponentially in the number of varied parameters ($x_i$).

This evaluation of the utilized building blocks underlines that they are suitable for use in our proposed designs. Their performance is satisfactory for set sizes and numbers of elements that we expect for real-world applications. Moreover, we show that our utilized building blocks largely scale linearly (only for the similarity metric, we observe an exponential blow-up). Consequently, in the following, we join these building blocks and study the performance of the entire designs in detail.

#### 5.2.2.3 Performance Evaluation of our Exchange Platform Designs

After establishing that the utilized building blocks scale to our needs, we now study the overall performance (**G-E4**) of our designs in more detail. To this end, we first evaluate them using synthetic input data. Afterward, we look into their performance when exchanging information for our real-world applications (cf. Section 5.2.1.2).

##### General Performance

We split our evaluation into workflows of data providers (Phase I: Data Provision) and clients (Phases II and III: Matching and Record Retrieval) because they are independent of each other during regular operation. Since our modular concepts only offer alternatives in Phases II and III, the data provision is identical for our designs BPE and PPE, i.e., we do not distinguish them when evaluating Phase I.

**Data Provision.** For the data-provision phase, we evaluate two crucial aspects: In addition to the impact of the number of records that a data provider processes, we also study the influence of the record length. First, in Figure 5.9, we detail the runtime when offloading up to 1000 records and show that the runtime scales linearly

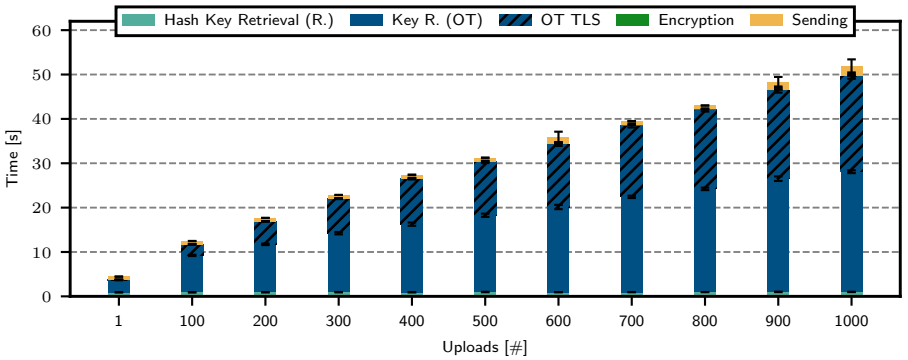



**Figure 5.9** As part of the data-provision phase, obliviously obtaining the encryption keys from the key server (using OTs) is the most time-consuming step when offloading data records.

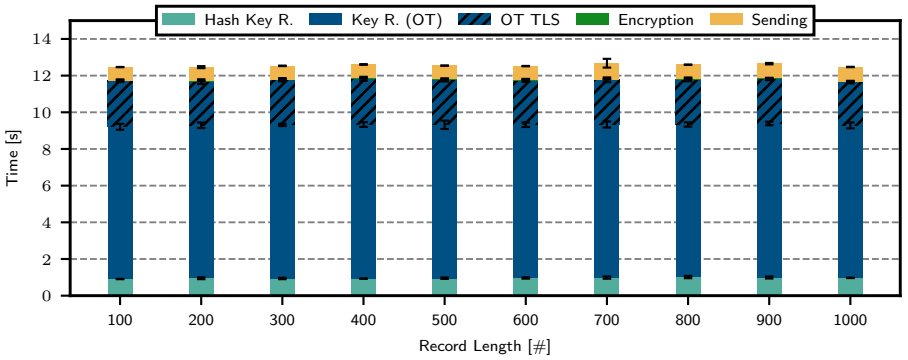



**Figure 5.10** The record size has a negligible impact on the runtime of the data provision.

with the number of records. We used records with 100 parameters, each representing a float uniformly chosen at random. Accordingly, all records have unique identifiers and require a unique key for encryption. The key retrieval dominates this phase.

Second, in Figure 5.10, we illustrate the influence of the record length on the data-provision performance. We uploaded 100 random records to the storage server for this measurement while varying the number of included parameters $m$. The key retrieval remains constant as the same (number of) encryption keys are retrieved for each measurement. The length of the records does not have a significant influence on the runtime as the key retrieval dominates the overall runtime of the data provision.

The observed performance of the data provisioning is reasonable, especially because it constitutes a one-time task for data providers (the exchange platform distributes the offloaded records to querying clients without requiring their involvement).

**Matching and Record Retrieval.** In contrast to the data provision, which constitutes a one-time task, client requests can be more time-critical. The real-world complexity of queries and the number of matches to be retrieved from the platform is highly use case-specific. Given that the performance of these phases depends on the utilized building blocks of our designs, we individually discuss BPE and PPE.

*BPE.* For this measurement, we offloaded records with 100 parameters and used 10 parameters as input for the indexing ($n = 10$, $m = 90$). Each input parameter

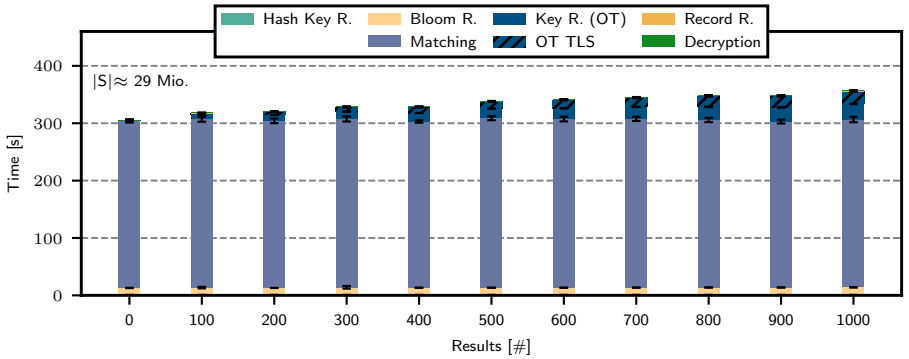



**Figure 5.11** Computing the similarity metric dominates the retrieval of data records.

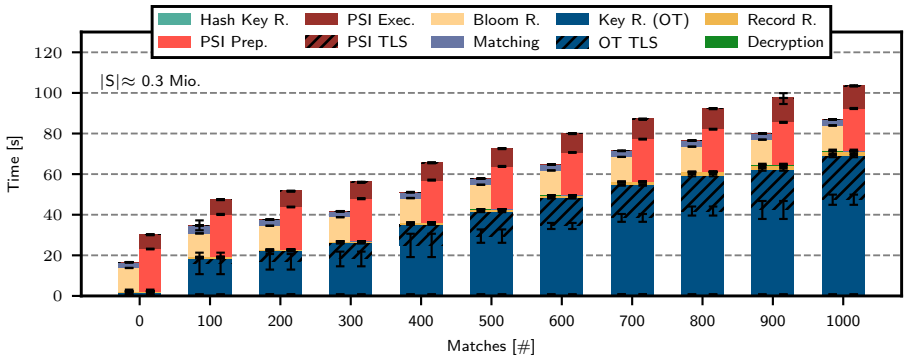



**Figure 5.12** If the set sizes are feasible to apply PPE, its performance is comparable to BPE.

is discretized to three digits. As similarity metric $s(q)$, we computed an offset of 0.5 % for each input record $q$. We ensured that sufficient records were matched and available on the storage server. While OTs mainly impact the provision, Figure 5.11 shows that the matching dominates the client queries. Despite the good performance of Bloom filters, the matching is expensive as it results in a large candidate set $S$ of >29 Mio. elements. While we observe a runtime below 5 min, real-world metrics might produce sets that are magnitudes larger, further increasing the total runtime.

*PPE.* We repeated the previous evaluation with a PPE-feasible-sized candidate set $S$ through a relative offset of 0.3 % with only 0.3 Mio. elements (compared to 29 Mio. elements). This limitation is not PSI-specific but follows from the memory constraints of our evaluation server, i.e., enterprise servers could support larger candidate sets. In Figure 5.12, we visualize the corresponding results. In comparison to BPE, the PSI-based matching in PPE introduces a slight overhead. Moreover, when comparing the different processing steps of our designs within the matching phase, we observe a comparable performance in BPE and PPE (Figure 5.13). In this setting, the primary overhead follows from the preparation of the PSI-based matching.

Overall, we notice that the runtime for client queries, involving both the matching and the record retrieval, is feasible for practical use because they may take up to several days. Consequently, our designs also support similarity metrics with (more) excessive candidate sets. In BPE, the testing for membership of indices does not even depend on external entities. Hence, this step is embarrassingly parallelizable, and clients can scale their metrics to their constraints, i.e., time and computational

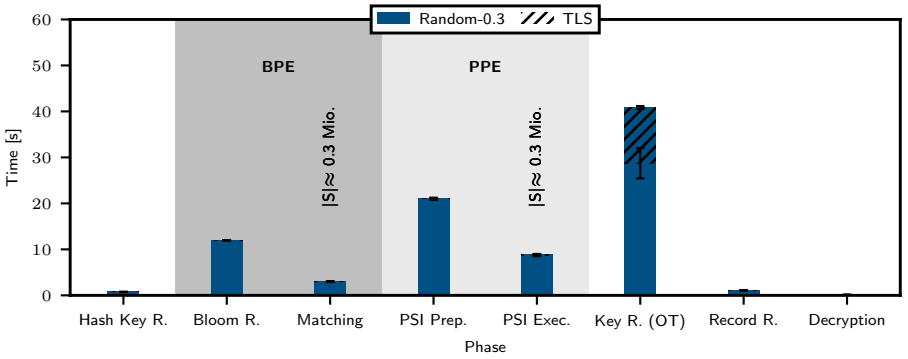



**Figure 5.13** For small candidate sets, the matching runtimes in BPE and PPE are comparable.

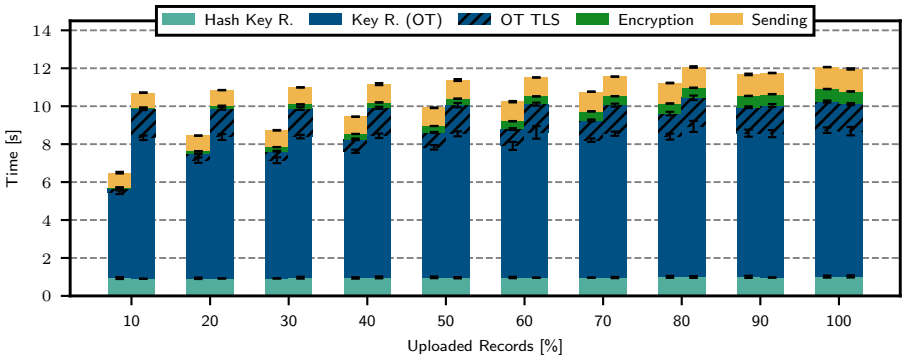



**Figure 5.14** Just like our general evaluation (Figures 5.9 and 5.10), our real-world application in injection molding also confirms that the key retrieval dominates the data-provision phase.

resources. In contrast, the matching in PPE requires sufficient resources at client and storage server, impairing its scalability. Accordingly companies can balance their confidentiality and performance needs in real-world deployments as needed.

### Performance Evaluation of our Injection Molding Application

To evaluate the applicability of our designs for the injection molding application (cf. Section 5.2.1.2), we now operate on a total of 4620 genuine records, consisting of 28 parameters each. These sensitive records describe the production of toy bricks: Each toy brick is defined by $m = 21$ geometry parameters, while the remaining $n = 7$ parameters describe 6 essential machine settings (injection volume flow, melt temperature, mold temperature, packing pressure, packing pressure time, cooling time) and one quality indicator (part weight). For other uses, apart from the optimization of machine settings during the process setup, the exchanged data and its representation will likely differ. Here, sensitive information is represented by the machine settings and the corresponding part quality: The data can only be used for transfer learning when being combined with the identifying parameters, i.e., geometry information With this indexing, we have a total of 60 indices, where each index points to 77 unique records that contain (slightly-)varying machine parameters.

**Data Provision.** As for our performance evaluation that utilized synthetic input data, we also study the data-provision phase for this application. In Figure 5.14, we

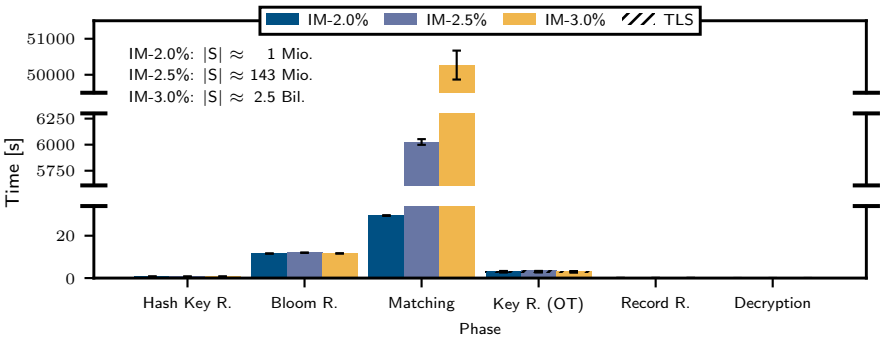



**Figure 5.15** The larger candidate sets produced by *IM-2.5%* and *IM-3.0%* lead to dominating matching phases for the client requests. This overhead only affects the querying client.

visualize the corresponding runtime for two scenarios: For the first measurement, shown by the left-sided bars, we chose to upload all parameters with the same identifier (index) before considering the data belonging to a different geometry. Here, the key retrieval overhead increases with a larger number of uploaded records because 77 records have the same index and therefore need the same encryption key. Therefore, the number of retrieved keys only increases when parameters with distinct identifiers are offloaded. For the second measurement, we selected the records uniform at random for each share. Here, already the first upload (10 %), equaling 462 records, has a high probability of containing one record of each of the 60 groups, such that all keys are required. Thus, the key retrieval times remain nearly constant over all runs. Since the entire provision phase takes less than 12 s, even if all records are uploaded at once, the provisioning is practical even in significantly larger settings.

**Matching and Record Retrieval.** For this application in injection molding, we evaluate three potential similarity metrics to study the diversity of client queries. For metric *IM-2.0%*, we used a relative offset for all (21) input parameters of 2 %. For *IM-2.5%* and *IM-3.0%*, the offsets are 2.5 % and 3.0 %, respectively. The rounding is set to two digits for each input parameter. In our example, both metrics result in a single match, i.e., the client retrieves the records that correspond to a single index.

*BPE.* In Figure 5.15, we visualize the different processing steps for both metrics. The matching step quickly dominates client queries, rendering the remaining processing steps negligible. We again underline that the locally-conducted membership tests are crucial in this regard. Keeping the application-induced time constraints (of several days) in mind, we can even support metrics with significantly-larger candidate sets with BPE. In conjunction with the virtually-irrelevant performance of offloading records, we thus conclude that the performance of BPE is well-suited for privacy-preservingly exchanging information. Consequently, BPE could greatly support transfer learning tasks in the domain of injection molding to ease the highly-complex and usually time-consuming task of setting up production lines.

*PPE.* For PPE, we can only evaluate *IM-2.0%* because the similarity metrics in *IM-2.5%* and *IM-3.0%* yield candidate sets that exceed the maximally-supported PSI set size of $2^{20}$. The evaluation of *IM-2.0%* with the PPE design (Figure 5.16) shows that the runtime of the PSI preparation step increases for larger candidate sets,

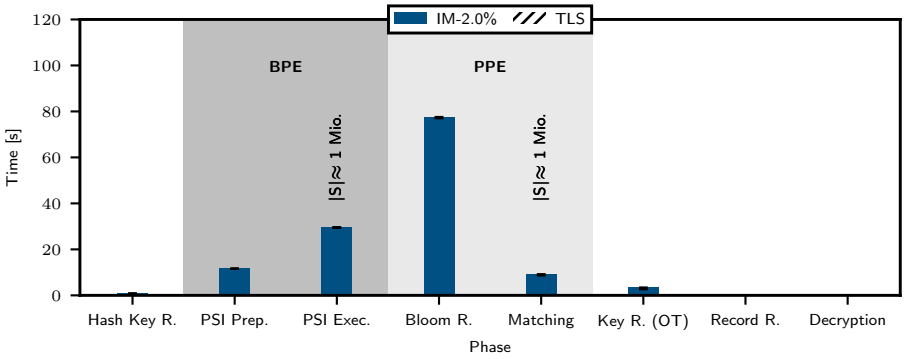



**Figure 5.16** While the PSI execution time stays constant (cf. Figure 5.13), the time to prepare the PSI-based matching increases with the size of $S$. Thus, BPE commonly outperforms PPE.

while the PSI execution takes approximately the same time. Moreover, these results imply that for larger candidate sets, the PSI-based matching adds significantly more overhead than the Bloom filter-based matching in BPE. This observation captures the primary drawback of PPE and explains why we favor the performance of BPE.

After this first evaluation of a real-world application, we present and discuss the performance of a second real-world application (cf. Section 5.2.1.2) in the following.

### Performance Evaluation of our Machine Tool Application

For our second application, which covers machine tools, we rely on a dataset with 600 records with 19 parameters each ($n = 17$, $m = 2$). Moreover, each record has a unique index. We again evaluate two independent client queries to investigate our designs' universality. First, for *MT-Material*, we only vary the production material of a workpiece, i.e., the client wants to produce an identical workpiece with another material. Second, for *MT-Diameter*, we request records where the same workpiece should be produced with a different milling cutter. To this end, we iterate over the parameter $x_i$, which defines the milling cutter's diameter. Given that we only vary a single parameter for each metric, the resulting candidate set (as well as the number of matches) is tiny compared to the previous application in injection molding.

**Data Provision.** Given that each record has a unique index, the OT-based key retrieval dominates the overall runtime. Still, the offloading of all records is completed within 30 s, and, therefore, this phase is uncritical for any real setting and participating data providers. Thus, we continue with the remaining phases.

**Matching and Record Retrieval.** We detail the processing times for BPE and PPE as well as both metrics in Figure 5.17. Even though this application does not introduce any hard deadlines, concluding the client query after less than 1 min is very performant for frequent use. Given that we only vary a single input parameter for each metric, the resulting candidate set is tiny compared to the metrics that we evaluated for the injection molding application. The small candidate sets ($|S| = 11$ for *MT-Material*, and $|S| = 701$ for *MT-Diameter*) reverse our previous observation that the PSI preparation dominates the runtime (cf. Figures 5.13 and 5.16). For

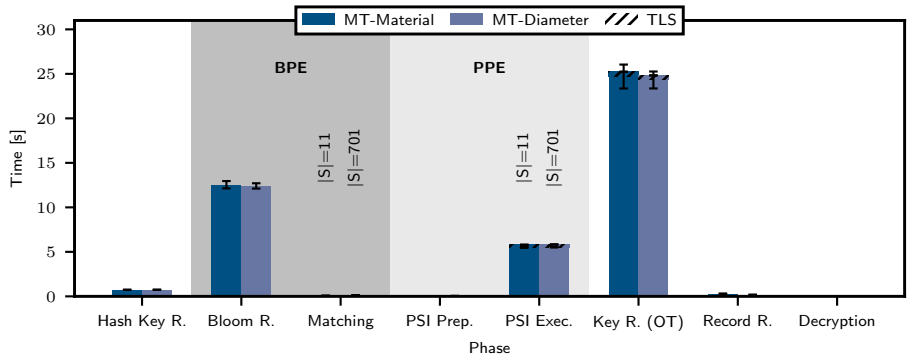



**Figure 5.17** In this smaller setting, the evaluated metrics produce negligible overhead for the matching. Hence, for these sizes and number of indices, both BPE and PPE are practical.

| | Client | Key Server | Storage Server | Data Provider |
|---|---|---|---|---|
| **G-E1: Provider Privacy** | | | | |
| BPE | ◕ | ●/○ | ●/○ | — |
| PPE | ● | ●/○ | ●/○ | — |
| OPE | ● | ●/○ | ●/○ | — |
| **G-E2: Client Privacy** | | | | |
| BPE | — | ● | ◕/◔ | ●/◔ |
| PPE | — | ● | ◕/◔ | ●/◔ |
| OPE | — | ● | ● | ● |

□/□: Malicious-but-Cautious Entities / Colluding Entities

**Table 5.5** We rate the achieved confidentiality guarantees per entity from ○ over ◔, ◑, and ◕ to ●. For provider privacy, we consider collusion among both servers, as it would allow for the decryption of all offloaded records. Likewise, for client privacy, we consider collusion among the storage server and a data provider. Jointly, they could then revert the hash identifiers and extract sensitive information about client queries. We omit the setting where a client colludes with both servers, as it would entirely violate provider privacy and reveal all records.

these settings, the PSI execution accounts for the primary runtime in PPE, and the overhead of preparing the PSI is negligible. The (large) Bloom filter and the key retrieval even dominate the total runtime, resulting in a situation where PPE outperforms BPE. The key retrieval times differ due to the number of matched and subsequently retrieved records (10 for *MT-Material* vs. 6 for *MT-Diameter*).

Using two real-world applications from two domains as well as synthetic input data, we have evaluated BPE and PPE in detail. We conclude that the overall performance is appropriate for practical applications and client queries. Hence, our designs are universally applicable and satisfy the outlined performance requirements (**G-E4**). In the following, we discuss their security and confidentiality guarantees to ensure that our exchange platforms are not only performant but also secure in practice.

#### 5.2.2.4 Security Guarantees of our Exchange Platform Designs

We now discuss the confidentiality guarantees of our designs. In particular, we study provider and client privacy (**G-E1** and **G-E2**). Without sufficient (technical)

guarantees, our designs are not realistic for use in the IIoT because companies are very cautious about sharing or exchanging sensitive information, especially across supply chains (cf. Section 5.2.1.1). In our discussion, we primarily consider malicious-but-cautious entities (cf. Section 2.1.2.1). By design, our exchange platforms do not require trust among (all) clients and (all) data providers, as each of them only interacts with the semi-trusted storage server. Thus, we individually discuss their interactions with the exchange platform and provide an overview in Table 5.5.

**Key Server.** As all sensitive key retrievals are handled via OT, the key server cannot harm provider and client privacy. While colluding with data providers does not harm the client privacy, colluding with the client could harm the provider privacy if ciphertexts have been retrieved illegitimately. Collusion of the operators of key and storage servers is the main threat in our designs as it would result in plaintext access, violating both provider and client privacy. However, a public entity operating the storage server (cf. Section 5.2.2.1) is unlikely to risk legal consequences following such misbehavior. Thus, we assume that the implications of this threat are limited.

**Storage Server.** During record retrieval, our designs do not hide the indices of client-requested records. Thus, the storage server can partially reconstruct the client's candidate set, slightly violating client privacy (**G-E2**). However, inferring the similarity metric is infeasible as neither the metric's input nor the unmatched indices are known to the server. Moreover, in BPE, handing out the Bloom filter affects the provider privacy as the client obtains an encoded representation of all indices with offloaded records. While Bloom filters do not support the listing of all stored items, brute-force attacks could still provide rough estimates, especially when configured with low false-positive rates. We tolerate this slight violation of provider privacy (**G-E1**) to enable the local computation of client metrics even in huge settings (up to billions of elements). PPE mitigates this slight violation of provider privacy because the PSI-based matching prevents clients from obtaining an overview of all filled indices.

By using a hash key for indices computation, which is unknown to the storage server, we increase both provider and client privacy as the storage server cannot compute any index itself even if it is aware of suitable input parameters. We achieve provider privacy as requested records are only shared without their origins against a payment. Similarly, providers are unaware of who paid for a record, satisfying client privacy. In the case of unintended data leaks, we protect records by utilizing different encryption keys to render brute-force attacks infeasible. Other misbehavior can be retraced through access logs at both key and storage servers. We leave an analysis of the privacy implications of joining these logs for future work.

To conclude, the confidentiality guarantees of our designs build on the separation of key material and ciphertexts. Our designs further require non-collusion among the storage server and data providers to ensure client privacy. Replacing the Bloom filter-based matching with a PSI-based matching slightly improves the provider privacy because clients do not have access to a probabilistic data structure that records all filled indices anymore. Likewise, relying on an OT-based record retrieval (as proposed in OPE) yields confidentiality improvements over the hash key-based retrieval (as used in BPE and PPE) because the storage server learns nothing about retrieved records anymore. We refer to Table 5.5 for a summary of these guarantees.

| | G-E1: Provider Privacy | G-E2: Client Privacy | G-E3: Deployability | G-E4: Performance | G-E5: Adaptability |
|---|---|---|---|---|---|
| BPE | ◕ | ◕ | ● | ◕ | ✓ |
| PPE | ● | ◕ | ● | ◑ | ✓ |
| OPE | ● | ● | ◕ | ○ | ✓ |

**Table 5.6** We assign scores from ○ over ◔, ◑, and ◕ to ● (✓) to describe to which extent our designs satisfy the design goals (cf. Section 5.2.1.3). Performance overheads impair the suitability for real-world deployments of our designs with stronger confidentiality guarantees.

#### 5.2.2.5 Comparing the Design Variants of our Exchange Platform

In this section, we have proposed a modular concept to realize a privacy-preserving exchange platform for the IIoT. Thereby, we can adapt the exact design according to use case-specific needs. In this context, we studied the performance implications and scalability of different building blocks, including Bloom filters, PSIs, and OTs. Looking at our designs, we observe that the runtime to offload records to the exchange platform as part of the data-provision phase is negligible, especially because it constitutes a one-time task for data providers. Our performance evaluations of synthetic inputs and two real-world applications (each with multiple queries) further demonstrate the feasibility of BPE. The reported performance even allows (i) clients to query and match large candidate sets $S$ in real-world settings (as required for queries in the injection molding application) and (ii) the exchange platform to handle a large number of indices. Thus, our designs are suitable for deployment.

By realizing the potentially-sensitive computations of similarity metrics $s(q)$ locally at the clients, we specifically address the required client privacy. While the Bloom filter-based matching in BPE slightly violates the desired provider privacy, the PSI-based matching in PPE reliably addresses this drawback. However, the computational overhead of PSIs mandates smaller set sizes, which limits the application areas of PPE. Our second application, which covers queries concerning machine tools, nicely demonstrates the practicality of PPE for real-world deployments.

We summarize this comparison as part of a design goal discussion in Table 5.6.

### 5.2.3 Takeaways and Future Research

In this section, we have presented and discussed our modular concept to realize exchange platforms for information sharing in the IIoT. With this contribution, we support the establishment of secure collaborations across supply chains because we address the stakeholders' confidentiality needs by providing technical guarantees. More specifically, our designs protect data providers, which offload their sensitive information, as well as clients and their sensitive queries. We now conclude our presentation by first discussing the suitability of our selected building blocks in Section 5.2.3.1. Afterward, in Section 5.2.3.2, we briefly highlight the universality of our proposed designs before outlining potential future work in Section 5.2.3.3.

#### 5.2.3.1 Suitability of Selected Technical Building Blocks

During conceptualization, two design aspects primarily shaped our selection of the utilized building blocks. First, we had to account for the sensitivity of the similarity metrics and the computation of matches, which is expressed by the goal of client privacy (**G-E2**). Accordingly, we could not design a protocol where the client offloads any of this sensitive information to other parties. Therefore, we had to create an approach that accounts for this need. Both of our matching concepts (locally-processed Bloom filters and PSIs) satisfy exactly this need. Second, following the decision to rely on two server components, we tried to reduce the trust in both servers to a minimum. By relying on OTs for the key retrieval, we reduce the trust in the key server to a minimum because it learns nothing about the exchanged information. As we have discussed in Section 5.2.2.1, an OT-based record retrieval (as utilized in OPE) does not commonly scale to secure collaborations in the IIoT. Hence, we consider a hash key-based record retrieval to be the second-best option for our designs.

Since we distribute our exchange platform over two operators (servers), we decided to not look into more complex building blocks, such as secret sharing. Regardless, with our modular concept, we support multiple building blocks by design. Consequently, we address the desired adaptability (**G-E5**) and allow for other realizations within the scope of our modular concept. Overall, we are confident that our selected building blocks are well-chosen since our evaluation generally attests a secure and performant operation of BPE and PPE. We attribute these findings to (i) the scalability properties of the utilized building blocks, (ii) the separation of data providers and clients, as well as (iii) the independent protocol sequence per company, i.e., a data-providing or querying entity does not depend on any other party (cf. **G-E3**).

#### 5.2.3.2 Universality of our Exchange Platform

When discussing the universality of our proposed designs that realize exchange platforms for information sharing across supply chains in the IIoT, we have to consider three dimensions: variation of supported applications, scalability of the designs, and diversity of the exchanged information. To assess the first dimension, we thoroughly evaluated two real-world applications (cf. Section 5.2.2.3), which allowed us to study application-specific variations in terms of scale (number of records, indices, and matches) and utilized similarity metrics in detail. Based on the results, we conclude that our designs are universally-applicable if we can trade off some confidentiality guarantees for improved performance. Likewise, concerning the expected scalability, our evaluation of the utilized building blocks and our general performance evaluation underline the feasibility of BPE and PPE. Moreover, given that most computational load is with the involved data providers and clients, we are confident that our concept ensures scalability-driven universality. Besides, we can easily scale out our exchange platform by distributing the index over multiple storage servers (or key servers). This way, our designs allow for massive indices that would otherwise exceed the supported set sizes in OTs or PSIs. Finally, regarding the diversity of exchanged information, we underline that our designs are capable of indexing and exchanging arbitrary records. Consequently, with an appropriate indexing scheme,

our designs support diverse information beyond process parameters, for example, to secure and facilitate exchanging best practices [PSF+23] among businesses.

#### 5.2.3.3 Future Work and Next Steps

While our evaluation underlines the secure and performant operation of our prosed exchange platforms, they are not yet ready for real-world use in the IIoT because the concrete design of the billing mechanism is still open. Hence, in preparation for real-world deployments, future work should add a (privacy-preserving) payment mechanism (cf. **G-E3**). In this regard, we envision a proxy-based approach, using either a trusted third party or a mix network [Cha81], to still reliably decouple data providers from clients and vice versa. In the long run, future work could look into how to rate the value of exchanged process data and, thereby, unlock new business models for data-providing companies. Such advances would greatly improve the acceptance of this type of secure collaboration while encouraging companies to provide their (sensitive) information. Similarly, researchers could look into ways to transform exchange platforms into a subscription model to ease the billing.

In a different direction, we identify three aspects that could further improve our designs and their usefulness. First, in their current realization, our designs do not provide strong accountability guarantees for all involved stakeholders. Accordingly, improving the auditability of offloaded and queried records could be researched to address respective needs. Second, concerning BPE, we could study the benefits of utilizing more advanced Bloom filters, such as multi-dimension Bloom filters [HX06, CL15], or applying differential privacy to the Bloom filter (cf. [EPK14, XVHS20]). Third, researching sophisticated approaches to come up with space-efficient indexing schemes would improve the processing of indexed records. Corresponding advances would be especially instrumental in easing the derivation of meaningful similarity metrics $s(q)$ for clients. However, our evaluation of real-world applications showed that even trivially-chosen identifying parameters ($x_i$) are indeed practical.

This subsection concludes the presentation of our fourth contribution, which supports secure collaborations across supply chains by providing a privacy-preserving exchange platform that is oblivious of both index records and client queries. With our work, we provide companies with a desired tool (cf. Section 5.2.1.1), address their confidentiality concerns, and improve information sharing across supply chains. Overall, in this chapter, we have tackled two crucial settings. In addition to proposing designs for privacy-preserving comparisons in industry (our third contribution), as part of our fourth contribution, we have developed a modular concept that enables privacy-preserving matchings in the IIoT. While the former is usually without direct implications on established processes, the latter is more precarious for the involved stakeholders because (i) matched and ultimately retrieved information is likely fed directly into running processes and (ii) the exchange of information also involves unknown, i.e., untrusted, companies. With these contributions and the general concept of secure collaborations in mind, in the next chapter, we analyze how far the industrial landscape has already evolved in recent times.

# 6

# Appraisal on Secure Industrial Collaborations

Our primary contributions (Chapters 4 and 5) underline that the conceptualization and realization of secure collaborations are practical from the information security dimensions. Building on this insight, in Section 6.1, we outline the readiness of the other dimensions (cf. Figure 1.2) for deploying secure collaborations in the (evolving) industrial landscape. Without sufficient advances in these dimensions, the global establishment of secure collaborations in the IIoT will not succeed. In this context, we further assess how our contributions impact the different dimensions as well as the ongoing evolution of the industrial landscape. Afterward, as part of an excursus, we utilize our experience in interdisciplinary cooperations to derive a research methodology that organizes interdisciplinary research efforts. We present the corresponding abstract process cycle underlying this methodology in Section 6.2. Our rationale is to ease interdisciplinary research activities for researchers and practitioners alike.

## 6.1 A Look at the Current State

To accurately rate the impact of our contributions, we have to evaluate our research in light of all relevant dimensions that enable industrial collaborations. Thus, in the following, we first outline the progress of the different enabling dimensions in Section 6.1.1. Thereby, we provide an overview of recent developments on broader prerequisites (beyond the information security dimension) for secure collaborations in the IIoT. Afterward, we consider general strategic research directions of the evolving industrial landscape: By discussing them in the context of secure collaborations, we highlight how (our) research on collaborations fits into the overall evolution of the IIoT. Subsequently, in Section 6.1.2, we revisit our contributions in light of our research question (cf. Section 1.2.2). We conclude that we are indeed able to realize

secure collaborations in the IIoT. Finally, in Section 6.1.3, we detail our view on how secure collaborations will further develop and evolve in the future.

### 6.1.1 Today's Situation in the Industrial Landscape

When recapitulating why we are observing an evolution of the industrial landscape at a significant pace, we quickly encounter the IoT. The momentum of the IoT, along with digitization and digitalization, reached the industrial landscape [Wit17, GKT19], which is the main driver of greatly-reshaped businesses and processes. The combination of these concepts coined the term IIoT, which expresses the digitized, interconnected, and networked state of companies in industry.

In this regard, we notice that means for ubiquitous communication and networking in the IIoT are widely available [RSS+16, WSJ17]. Likewise, sufficiently performant approaches that enable the processing and analyzing of vast amounts of data have matured toward practical use [uRYS+19]. However, despite this convincing state, isolated information silos still largely hinder the global exchange and dissemination of knowledge and information, as we have outlined in Section 1.1. As a result, the desire for collaborations in the IIoT emerges. But, their acceptance depends on realizing them securely. Fortunately, as we have pointed out in Section 2.4, sufficiently many and diverse building blocks to secure them are conceptually available. Consequently, now is the time to research how to widely and securely establish collaborations in the IIoT. To contribute an essential aspect and to create a profound foundation for practical deployments, in this dissertation, we have successfully worked on this research gap by tackling the realization of secure collaborations from the information security dimension, which has particularly hindered their real-world use so far.

#### Progress in the Enabling Dimensions of Industrial Collaborations

As we have established in Section 1.2.1, the successful establishment of secure collaborations in the IIoT depends on multiple dimensions (cf. Figure 1.2). In addition to the information security dimension, which was the focus of this dissertation, the economic, legal, operational security, and interoperability dimensions are critical to pursuing the desire of having collaborations in the IIoT. Since we revisit our contributions in light of our research question in Section 6.1.2 anyway, we also defer the assessment of the information security dimension. Hence, in the following, we illustrate how research in the other dimensions has prepared the industrial landscape for secure collaborations so far and which individual research challenges remain.

**Economic.** Most importantly, economic questions are critical for stakeholders to assess whether they are willing to participate in industrial collaborations [Hor01, Sah03]. Moreover, establishing new or improved business models to transform the industrial landscape is an important, open factor in encouraging real-world deployments that utilize secure collaborations [MBP23]. Especially the global dependencies and implications when striving for softer factors, such as ESG goals [FBB15], urgently mandate that stakeholders increasingly consider the consequences of collaborations as part of their strategic organization [KJPE23]. Apart from the primary

focus on business models, additional research on (i) how companies behave in light of secure collaborations (e.g., studying game theory [ZPMG19]) and (ii) widely-applicable approaches to deterministically rate the value of information is needed, with the latter recently gaining interdisciplinary interest and first interesting results [TBP23].

**Legal.** Since deployments and realizations of secure collaborations are rare and mostly vague, legislative ramifications are also unclear, especially when considering liability aspects. In the long run, elaborate legal frameworks need to establish accountability from a legislative perspective on top of the technical accountability guarantees [PMK+21]. Currently, the legal implications of deeply-rooted secure collaborations are not yet well understood. Specifically, how to deal with governmental oversight in collaborations with confidentiality guarantees is a critical aspect. Likewise, regulating access permissions and transparency is important, i.e., handling the platform openness [PMK+21]. Otherwise, unfair competition or misbehavior that follow from secure collaborations are hard to detect. Finally, research still needs to ensure compliance with legislation [PDG+19] and publish guidelines for stakeholders in the IIoT. In this regard, ownership responsibilities, liability concerns, and law-abiding processing of usage information from consumers (cf. GDPR [KPW21]) increasingly gain relevance with the continued emergence of secure collaborations.

**Operational Security.** In this dimension, reducing attack vectors in a converged IIoT that also follows from the establishment of secure collaborations is an important research aspect, which also entails direct implications for the safe operation of industrial environments. Looking into this direction, measurement studies [DLF+20, DLP+22, DSD+23, DHL+24] outline that stakeholders in the IIoT have to catch up with security best practices to secure their operations. However, the technology, attested concepts, and protocols to do so are readily available. We refer to related work [SHH+21], which proposes five groups of measures to secure operations in the IIoT. Especially the potential of industrial intrusion detection is significant because PPC and running processes, as well as production-to-production communication (cf. Section 2.1.1), express distinct patterns while revealing correlations between physical observations and communication, which can be exploited for intrusion detection. The multitude of research in this field underlines the relevance of this direction [KWP+22a, KWP+22b, WTvS+22, WWSH22, WKW+23, LWW+23].

**Interoperability.** Finally, to realize interoperability across stakeholders, consistent identifiers for information and entities that are globally usable are highly desirable. Moreover, standardization should strive for unobstructed yet secure collaborations by ensuring platform openness. Here, we specifically look forward to the ongoing developments of the outlined large-scale initiatives (cf. Section 1.1) that are closely cooperating with major companies and cloud service providers. The FAIR data principles can support these developments because they are not limited to use in the context of research data anymore: Related work increasingly adapts them for the IIoT, e.g., FactDAG [GPL+20, GPT+21, Gle23] or FAIR sensor services [BMS21, BPM+23]. In addition to technical interoperability, research further considers the suitability of human-machine interfaces [BSZ22], i.e., the interoperability between humans and machines. Without appropriate approaches, the workforce in the industrial landscape will not be able to configure secure collaborations or fully source their benefits.

Overall, we conclude that research activities for most dimensions (except for the legal dimension) are ongoing and generally well on track, i.e., research braces itself for the establishment of secure collaborations and their corresponding needs. In particular, research proposes new concepts where needed and builds on established technologies where possible. Hence, these developments contribute to the prerequisite of realizing widely-accepted, successful, and impactful collaborations in the IIoT. Given that the overall evolution of the industrial landscape is still in its infancy, we next discuss corresponding general research directions in the context of secure collaborations.

### Strategic Research Directions to Sustainably Evolve the Industrial Landscape

To assess the general (research) developments that follow from the desire to establish secure collaborations in the IIoT, we now look into relevant research directions that cover the overall evolution of the industrial landscape. In particular, we source an ongoing initiative [BDJ+22], which originates from the Cluster of Excellence "IoP" (cf. Section 2.1.1). While the general goals of the research initiative and this dissertation largely align, we still have to discuss their five strategic research directions specifically in light of secure industrial collaborations.

*Standardized (Data) Interfaces.* This direction captures aspects related to ensuring compatibility between collaborators and information flows, with a primary emphasis on interfaces. We place corresponding research needs in the interoperability dimension. Consequently, even though standardized interfaces would simplify the application of secure collaborations in practice, we largely consider corresponding research in this direction to be out of scope for the information security dimension.

*Interconnecting (Domain) Knowledge.* Globally connecting knowledge and information is a fundamental aspect for the evolved industrial landscape. To achieve this goal, the authors propose advances that incorporate autonomous agents and appropriate metadata. This goal is crucial to deeply integrate secure collaborations into the industrial landscape. In particular, the exact realization will have a profound impact on the topology of the IIoT and the chosen mode of operation (cf. Section 2.1.3).

*Burden-Free Operation.* Research in this direction focuses on reducing the overhead of participating in the IIoT. Consequently, it is essential for industrial collaborations as well and generally affects all dimensions equally because every dimension introduces its own hurdles and requirements for the operation. Research should specifically focus on proposing approaches that consider the concept of *usable security* [GL14] to stimulate the deployment and ease the application of secure collaborations. Naturally, secure collaborations, their integration, and automation will mature and evolve over time while possibly incorporating the findings of this dissertation.

*Real-World Integration.* Research related to this direction is crucial to truly demonstrate the benefits and impact of the IIoT, as well as of secure collaborations, on industry, consumers, and society alike. Since we focus on conducting fundamental research to assess whether we can realize secure collaborations, this research direction is the next logical next step following this dissertation, i.e., pushing our designs to practical use in the industrial landscape. We look forward to seeing how global knowledge and information sharing will tear down today's isolated information silos.

*Long-Term (Information) Usage.* At last, this aspect is tightly interwoven with the research that we presented in this dissertation. With secure collaborations, we support this research direction by proposing first approaches that provide stakeholders with technical guarantees of confidentiality, reliability, and accountability. Upcoming interdisciplinary research can ideally capitalize on these advances to further evolve the industrial landscape. After our initial steps, significant effort is still needed to push the collaboration-induced improvements to a multitude of diverse use cases.

After dissecting relevant strategic research directions in light of our focus on secure collaborations, we notice that the research conducted as part of this dissertation aligns well with the overall goals of initiatives that attempt to evolve the IIoT. In this regard, we have identified profound synergies that will help to push secure collaborations and substantial information sharing in the IIoT into practical use.

#### The Path Forward

In this dissertation, we have studied and evaluated real-world applications that utilize secure collaborations. The respective results already demonstrate the technical readiness of today's building blocks to reliably secure collaborations in the IIoT. In the next subsection, we will thus specifically discuss our findings in more detail.

Concerning the path forward, we want to emphasize at this point that additional real-world applications that build on (and benefit from) secure collaborations will generate more acceptance, which will bring more ideas concerning novel collaborations and how to evolve established information flows. This evolution will then bring more applications to the IIoT and so forth, effectively exhibiting a mutually-reinforcing evolutionary process. Combining this development with a continuous review of the situation in the industrial landscape will provide a better understanding of collaborations, their security, suitable technical concepts, realized benefits, and introduced risks for all involved stakeholders, practitioners, and researchers.

### 6.1.2 Research Impact of this Dissertation

The goal of this dissertation is to contribute to paving the way for secure and reliable information sharing in the IIoT by establishing secure collaborations. Accordingly, we discuss the presented findings in light of our primary research question in the following. Subsequently, we summarize our contributions to establishing secure collaborations in the IIoT before assessing the overall impact of our work.

#### Quo Vadis: Can we already realize Secure Industrial Collaborations?

Our primary research question, which we originally introduced in Section 1.2.2, reads as follows: *How can we enable secure industrial collaborations in real-world settings?*

As a foundation for our contributions, we conducted a survey of promising technical and conceptual building blocks in Section 2.4. Our findings highlight that we

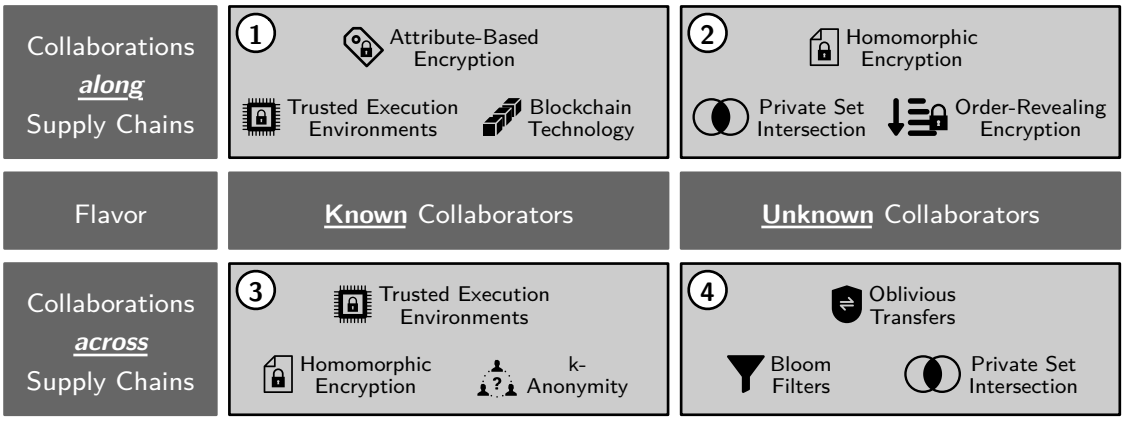



**Figure 6.1** As part of our contributions (①–④) to secure collaborations, we utilize various technical building blocks. We selected them according to the needs of the addressed use case.

have a wealth of different concepts and approaches available that promise to facilitate the realization and establishment of secure and reliable collaborations in the IIoT. Sourcing this result, we selected various technical building blocks according to the needs of the tackled collaboration, as we illustrate in Figure 6.1. With this overview of applicable and practical building blocks, we contribute to answering the subquestion of which technical means should be sourced for collaborations in the IIoT. Unfortunately, as expected (cf. Table 2.1), no one-fits-all solution exists. However, in terms of the information security dimension, our designs underline the suitability of private computing to securely realize collaborations in the IIoT. Since we evaluate the designs of every contribution using two real-world applications each, we further make sure that we do not overfit them to a specific application, boosting our confidence that they are indeed widely applicable in real-world settings.

Due to our technical focus on the information security dimension (cf. Section 1.2.1), we can only speculate whether our designs are suitable to convince stakeholders to participate in industrial collaborations. However, given that we generally address the (strict) confidentiality goals of each setting and use case, involved stakeholders would likely need to cite other arguments as to why they are not willing to participate in appropriately-secured collaborations. Accordingly, we table this subquestion until the first secure collaborations have been deployed for commercial use.

While this dissertation sheds a lot of positive light on industrial collaborations and the potential to realize them securely, we also identify two major (technical) obstacles that have hindered collaborations across supply chains so far. First, providing accountability guarantees is already challenging when dealing with untrusted collaborators on its own. Utilizing building blocks to realize privacy-preserving information flows further exacerbates this situation because research on verifiable computing is still in its infancy, and applying building blocks that entail accountability guarantees partly contradicts the confidentiality needs of stakeholders in the IIoT. Second, although concepts and building blocks to realize comparisons, matching, and information sharing and retrieval are readily available (as we have shown in this dissertation), use cases in the IIoT significantly challenge their large-scale applicability because they are hardly performant enough to scale to industry needs. Consequently, as we

noticed while designing our approaches, optimizing and tweaking protocols to specific use cases is crucial to ensure their practicality. However, this challenging aspect is time-consuming and depends on tightly-intertwined interdisciplinary research.

Our dissertation stresses that we are able to reliably secure collaborations that build on practical, real-world applications. Since we focused on the fundamental research challenges of secure collaborations while facing real-world-induced constraints, i.e., we conducted use-inspired basic research [Sto97] in this dissertation, we cannot assess whether our findings would truly convince stakeholders in the IIoT to rely on them. However, from a factual perspective, we are convinced that they will.

### Scenario-Focused Summary of our Contributions

While pursuing our research question, we worked on four primary contributions that tackle secure collaborations along and across supply chains in the IIoT. In the following, we highlight the most important takeaways for the big picture of establishing and advancing secure collaborations and refer to Chapters 4 to 5 for elaborate discussions on the respective technical foundations and evaluation results.

**A Processing Pipeline for Reliable Information.** Our first contribution establishes end-to-end security for arbitrary information flows along supply chains. That is, we demonstrate the benefits of trusted sensors and confidential computing for the reliability and authenticity of information that originates from (untrusted) remote environments. Sourcing technical building blocks, we further introduce long-term accountability guarantees to information that is processed along supply chains while supporting fine-granular access control with our blockchain-backed design PrivAccI-Chain. To implement practical access control and ensure confidentiality guarantees, we propose an efficient multi-layered encryption scheme that allows companies to trade off confidentiality and transparency, even in large-scale supply chain networks. Altogether, we reliably secure the processing of information from sensor to storage, even in scenarios where information is only shared or retrieved after a long time.

**Finding New Suppliers with Privacy-Preserving Purchase Inquiries.** For our second contribution along supply chains, we targeted a setting where collaborators are mostly untrusted, i.e., corresponding designs need to account for this challenge. Using well-established building blocks from privacy-preserving computing, namely, PSI and HE, we reliably protect companies' sensitive information during the early stages of procurement. More specifically, with our privacy-preserving purchase inquiries, we limit the information that they have to disclose upfront to a minimum. To the best of our knowledge, we are the first to emphatically express and appropriately address these privacy issues. Our designs, PPI, HPI, and cHPI, neatly integrate into established procurement processes. Consequently, we enable companies to consider a larger number of potential buyers and sellers without fearing for their competitive advantage, even when primarily dealing with untrusted collaborators.

**Privacy-Preserving Company Benchmarking.** Concerning secure collaborations across supply chains, we have identified the shortcoming of insufficiently-protected yet valuable benchmarking algorithms. Accordingly, our third contribution covers

company benchmarking. For this type of collaboration across supply chains (without direct implications on running processes), we proposed, implemented, and evaluated two approaches, HW-PCB and SW-PCB, which build on fundamentally-different concepts from private computing (TEEs and HE). Our representative real-world evaluations show that these concepts are practical and sufficiently scalable for deployment in the IIoT and demonstrate that secure collaborations across supply chains can indeed succeed. That is, by showcasing the readiness of technical building blocks for use in the IIoT, we create the first step for many more collaborations in industry.

**Privacy-Preserving Parameter Exchange.** As our fourth contribution, we focus on establishing a new type of collaboration in the industrial landscape. Specifically, by proposing a modular concept that promises privacy-preserving exchange platforms for use in the IIoT, we address a desire by practitioners. Our designs, BPE and PPE, efficiently combine different building blocks, including Bloom filters, PSIs, and OTs, to reliably and efficiently secure the exchange of sensitive information. Our modular concept further allows trading off performance and confidentiality guarantees. With this work, we enable the utilization of knowledge across supply chains, even if unknown collaborators are involved, and further demonstrate the suitability and universality of our designs by evaluating two real-world applications. The resulting collaborations are essential to effectively convince stakeholders that the oblivious processing of sensitive information in the IIoT is attainable, i.e., research succeeds in securing collaborations while providing sufficient guarantees that allow stakeholders to directly adapt running processes based on received information.

Altogether our contributions provide solid insights into challenges, requirements, and best practices when realizing secure collaborations in the IIoT, both along and across supply chains. Most importantly, we highlight that securing collaborations in the IIoT while introducing technical guarantees is feasible (and scalable) with today's well-established building blocks. Large-scale initiatives and future research can capitalize on our findings when proposing and evolving secure collaborations.

### Lasting Impact of our Contributions

To expand beyond the immediate benefits of this dissertation, we now consider the lasting impact of our work. While we can only speculate whether our contributions will have a lasting impact, several observations increase our confidence in the (lasting) value of our findings, as we further outline in the following.

**A Processing Pipeline for Reliable Information.** Despite numerous research in the past, we still advance the most basic form of collaborations in the IIoT by providing means for sophisticated information sharing and utilization. In particular, we firmly stress the challenge of reliable sensing, which research is barely addressing at the moment (and has also largely been overlooked in the past). With the proposed use of trusted sensors, we set the stage for various advances in research that provide technical guarantees in this largely-overlooked area. Specifically, we want to point out (i) the potential benefits of sourcing technical benefits as part of reputation systems and (ii) advancing research into correlation and threshold-based sensor manipulation following the technical reliability guarantees E2E sensing is introducing.

**Finding New Suppliers with Privacy-Preserving Purchase Inquiries.** For an entirely different collaboration setting, we create the foundation for a new research direction that particularly focuses on the confidentiality needs of stakeholders during the early stages of the procurement process. Our contribution is only the first step to fully addressing the research gap. An independent review board by the German Federation of Industrial Research Associations [AiF98] and the FQS Research Association [FQS01] rated our initial suggestions as a promising research proposal due to their relevance and innovativeness for companies in the industrial landscape. Consequently, this dissertation is not the end of the road for corresponding research.

**Privacy-Preserving Company Benchmarking and Privacy-Preserving Parameter Exchange.** Since collaborations across supply chains have rarely been established in practice so far, we jointly discuss our corresponding contributions. Most importantly, we have convincingly demonstrated the feasibility of secure collaborations across supply chains. Therefore, we are confident that our contributions are well-suited to convince stakeholders to be open to information sharing across supply chains, potentially even with competitors. We substantiate our findings by demonstrating respective collaborations in a spectrum of situations, from privacy-preserving comparisons that have no direct impact on running processes to privacy-preserving matchings that are more likely to influence (local) processes based on globally-sourced knowledge. To the best of our knowledge, in contrast to related work and large-scale initiatives that largely build on organizational trust guarantees [LPMW22], we are the first to primarily rely on technical (confidentiality) guarantees for several challenging use cases in the IIoT. Thus, we add a valuable (and currently missing) perspective to research and the ongoing evolution of the industrial landscape.

Leaving our individual contribution aside, today's (research) visions, such as the IDS (cf. Section 1.1) or IoP (cf. Section 2.1), cannot convincingly succeed without secure industrial collaborations, mostly because they generally depend on tapping into today's isolated information silos. Hence, the research directions that we addressed with our contributions and associated considerations of security, privacy, reliability, and accountability are crucial for their lasting success as well as for the acceptance by stakeholders in the IIoT. Consequently, we anticipate that this dissertation and its findings will also impact their future developments.

With this discussion, we conclude our presentation of this dissertation's research impact and proceed to an outlook on aspects that likely influence secure collaborations, specifically in the information security dimension (cf. Figure 1.2), in the future.

### 6.1.3 Outlook: Advancing Secure Collaborations

After outlining the impact of our contributions, we now discuss in which conceptual directions secure collaborations might evolve. Afterward, we further point out which additional technologies from computer science could become increasingly important for secure collaborations due to their value for the global utilization of knowledge. Finally, we sketch the path forward for the most relevant technical building blocks from private computing, which we also already utilized in our proposed designs.

### Future Developments of Secure Collaborations

Apart from highly-specialized and use case-oriented collaborations, which we disregard in this general overview, we identify three main strains of future developments for secure collaborations in the IIoT [PMK+21]. While the first strain on *reliable product information* is primarily relevant for all sorts of collaborations along supply chains, the other two strains cover collaborations across supply chains, with *efficient and dependable collaborations* improving the information sharing with known collaborators and *distributed data markets* focusing on unknown collaborators.

*Reliable Product Information.* The first strain goes beyond what we tackled in our first contribution (our processing pipeline) as it covers all kinds of developments related to digitally-attested information in the industrial landscape. Transforming the industrial landscape takes time, even if suitable approaches are readily available. Hence, process and operation adjustments that follow from improved reliability and accountability (e.g., following digital transmission contracts [MGP+21]) still have to emerge, i.e., the end of the evolution concerning reliable information has not yet been reached. As we have outlined before (cf. Section 6.1.1), the interplay with the legal dimension will greatly impact their evolution. In addition to the focus on accountability, this strain further covers (technical) approaches that improve the linking of physical and digital goods because respective advances will also affect if and how secure collaborations can provide which kind of guarantee (e.g., stronger accountability guarantees). Overall, this strain of research supports the establishment of more substantial secure collaborations of companies along supply chains.

*Efficient and Dependable Collaborations.* Since collaborations across supply chains are barely utilized at this point, corresponding advances promise significant potential. The usefulness of these collaborations increases significantly if companies can depend on the information they are receiving. Respective developments go hand in hand with new business models and approaches, such as MaaS or SCMaaS, which increasingly shape (secure) collaborations in the IIoT. In this context, information sharing across supply chains is essential to acquire a broad knowledge base for these services. In a similar direction, companies are likely to increasingly adapt their operation based on the entire product lifecycle, i.e., production, development, and usage [PGH+19], when making (business) decisions. Hence, with truly dependable collaborations, companies can globally source knowledge on how to process respective information and retrieve best practices on how to react to specific findings.

*Distributed Data Markets.* Finally, as we have pointed out as part of our fourth contribution (our exchange platform), properly and widely incorporating the value of information and knowledge into collaborations and the IIoT will have a significant impact on the industrial landscape and its evolution. Since data is being referred to as the new oil (cf. Section 1.2.1), the emergence of data markets in the IIoT and their integration into secure collaborations both along and across supply chains seems to only be a question of time. So far, research has primarily considered them in the context of personal data [MMZ+17]. However, we expect that upcoming research will reliably transfer this concept for application in the IIoT. Essentially, the developments of large-scale initiatives (cf. Section 1.1) and their pursued concept

of data ecosystems [GVC+22] match the idea of sophisticated data markets (under a newly-coined term). However, by using a different term for this strain of research, we emphasize that technical guarantees (as prevalent in secure collaborations) are mission-critical to successfully establish data ecosystems in the industrial landscape.

After highlighting the conceptual strains of development, we now shift our discussion to technologies that promise to extensively evolve secure collaborations in the future.

### Technologies to Further Evolve Secure Collaborations

Apart from the building blocks that we have surveyed before (cf. Section 2.4), we see two (emerging) technologies as very promising tools to extensively transform the scope and application areas of secure collaborations in the future.

*Federated Learning.* First, when reflecting on our contributions to secure collaborations across supply chains, we notice that their invasiveness for the involved stakeholders is increasing. Pushing this observation forward, we would move from privacy-preserving comparisons over privacy-preserving matchings to privacy-preserving machine learning [PHW21]. In particular, the widespread utilization of federated learning [LSTS20], a privacy-improving distributed concept in the domain of machine learning, is a promising candidate to combine the knowledge of multiple stakeholders in the industrial landscape, i.e., it promises to tear down today's information silos. But first, to enable the use of federated learning for secure collaborations, research must fully comprehend the associated privacy risks. Additionally, privacy-preserving machine learning offers the possibility to automatically adapt running processes. However, the implications of such invasive approaches are not yet fully captured and understood. For example, how reliable are the achieved results, how can we verify their correctness, and which entity is liable for any damages. If research does not address these issues, uncertainties concerning the technology will prevent the application of federated learning in settings with cautious stakeholders.

*Process Mining.* Second, linking the information from various stakeholders in the IIoT allows companies to comprehensively study their processes and operations in their entirety, both along and across supply chains. In this context, federated process mining [vdA21] is essential to account for the confidentiality needs of involved stakeholders. Moreover, the wealth of processes and corresponding information on them allows for in-depth analyses of how they interact with each other and how the overall industrial landscape is performing. Related work coins this development as object-oriented process mining [vdA19]. We look forward to seeing the impact of IIoT-compatible process mining on secure collaborations once this technology has fully matured to account for the confidentiality needs of stakeholders in the IIoT.

While the potential of these two technologies is substantial to extensively affect secure collaborations in the future (they promise significant usefulness improvements by combining globally available knowledge), our surveyed building blocks, which we also utilized in our designs, will also advance over time. In the following, we thus discuss the road ahead for two prominent building blocks of private computing.

### Growing Importance of Today's Building Blocks

To reliably secure collaborations in the IIoT, building blocks from the area of private computing promise great potential because they (i) are able to provide technical guarantees, e.g., on confidentiality, and (ii) they are well-established concepts. As part of our contributions, we have already capitalized on these properties.

We are optimistic about seeing significant advances and future developments for building blocks from private computing because large communities constantly work on improving them, and research has recently seen several improvements [VJH21]. In the area of privacy-preserving computation, `homomorphicencryption.org` [hom17], `FHE.org` [FHE21], and OpenFHE [Ope22] focus on increasing the application areas of HE through standardization efforts, by supporting additional tasks, or by designing more scalable primitives. Likewise, the Confidential Computing Consortium [The19] drives standardization efforts in the area of confidential computing to improve adoption and market penetration. Research on secure collaborations can only benefit from advances in the area of private computing because the respective building blocks are essential to reliably secure the exchange of information in the IIoT. Especially additional awareness and publicity can help (i) to broaden the research in this area and (ii) to convince conservative stakeholders of their usefulness, security, and performance. Even if these building blocks are primarily researched for and applied in other settings (e.g., cloud computing), matured and evolved concepts still indirectly benefit research and, eventually, real-world deployments in the IIoT.

While the traction of these highlighted communities is already substantial today, we expect that the importance of research on verifiable computing (as recently exemplified through fundamental developments [FNP20, VKH23]) will gradually increase over time because corresponding building blocks are able to holistically provide further technical (confidentiality) guarantees on privacy-preserving computations. Apart from applications in blockchain technology, cloud computing, or electronic voting, corresponding advances would also be beneficial to evolve secure collaborations in the IIoT, as long as the building blocks scale to industry-sized applications.

At this point, we again refer to the mutually-reinforcing evolutionary process of enabling building blocks and secured use cases (cf. Section 6.1.1) because advancing the building blocks can contribute to the realization of novel or highly-sophisticated secure collaborations. Simultaneously, the need for specific building blocks, e.g., as required for a specific use case or collaboration, can also trigger corresponding research efforts and initiatives. Thus, we look forward to the improvements that follow from the growing importance of building blocks from private computing.

This subsection concludes our appraisal on the current state of the evolution of the industrial landscape. We look forward to seeing how secure collaborations develop in the future and how they transform information sharing in the IIoT. Since the design and realization of secure collaborations are as important as the initial idea to look into them for a specific use case, in the following excursus, we pass on our gathered experience on how to conduct interdisciplinary research that involves computer scientists and practitioners alike by formalizing an abstract research methodology.

## 6.2 Excursus: Conceptualized Research Methodology

As we have motivated in Section 1.3, the evolution of collaborations in the IIoT depends on tightly interwoven interdisciplinary research. The development of novel, more sophisticated, and complex industrial collaborations will only succeed after establishing simpler and widely-relied-upon ones. Accordingly, research would benefit from best practices that guide and support corresponding efforts. In the following, we capitalize on experience that we collected while working on this dissertation's contributions and propose a conceptualized research methodology for interdisciplinary research at the intersection of information security and use cases in the industrial landscape. First, in Section 6.2.1, we elaborate on our motivation, summarize our contributions, and refer to other developments in the context of research artifacts and reusability. Subsequently, in Section 6.2.2, we detail our derived methodology with which we intend to improve tomorrow's interdisciplinary research cooperations.

### 6.2.1 Rationale Behind this Excursus

In this dissertation, given its importance, we focused on research in the information security dimension (cf. Section 1.2.1). However, from a more general perspective, the benefits of cybersecurity research (including operational security) are crucial for the success and safety of real-world applications. When looking at the IIoT, we identified great potential from applying methods and tools developed by the security and privacy community to industrial use cases as they might be able to provide functionality that was previously considered impossible due to prevalent confidentiality and privacy concerns. Corresponding advances require intensive cooperation among practitioners and cybersecurity experts to come up with suitable use case-fitting solutions. A sustainable evolution of the IIoT can only be achieved by combining security research with novel industrial applications. Unfortunately, as we have also experienced in the context of this dissertation, interdisciplinary research is challenging, and thus, lots of potential remains untapped so far. The traditional foci of both groups further intensify this issue. On the one hand, practitioners are more likely to be reluctant to share their information due to confidentiality concerns or cannot imagine the possibilities that use case-tailored privacy-preserving building blocks enable reliably. On the other hand, security experts might lack a sufficient understanding of industrial processes, a respective vision of future applications, and the required contacts to significantly advance real-world applications.

To mitigate this obstacle, we abstracted our interdisciplinary research experience into an abstract methodology that expresses the needs of such applied research. As we detail in the following, we also enrich our presentation with concrete and relatable examples to make the methodology accessible for experts from both domains.

#### 6.2.1.1 Our Contributions to Real-World Use Cases

Based on our experience while conducting interdisciplinary research in the IIoT, we derived a process cycle that methodologically describes such interdisciplinary

research. Primarily, we refer to the contributions (Chapters 4 and 5) and the use cases (Chapter 3) of this dissertation. We abbreviate these individual efforts as:

- *A Processing Pipeline for Reliable Information* **(ALONG-Pipeline)**,
- *Finding New Suppliers with Privacy-Preserving Purchase Inquiries* **(ALONG-Finding)**,
- *Privacy-Preserving Company Benchmarking* **(ACROSS-Comparison)**, and
- *Privacy-Preserving Parameter Exchange* **(ACROSS-Matching)**.

Wherever reasonable, we further refer to practical experience from (interdisciplinary) past and ongoing research activities within the Cluster of Excellence "Internet of Production (IoP)". Thereby, we augment our presentation, provide additional depth, and underline the universality of our proposed methodology. To illustrate the practical application of this methodology, we explicitly discussed our **ACROSS-Matching** contribution as a case study in our previous work [PBD+21]. To provide additional context to our work, we take a look at related work in the following.

#### 6.2.1.2 Related Work: Revisiting Artifacts and Methodologies

In computer science, research usually tries to come up with universally applicable approaches by deriving abstract models. This best practice intends to foster the reusability of work while reducing the amount of redundant research. Artifacts are being open-sourced and shared more commonly these days to reduce the effort needed in reproducing past work [BBE+22]. As a means to support open-sourcing, artifact badging [ACM20] might motivate researchers to share data and source code artifacts along with their papers, as papers with artifacts seem to receive a broader reception [ZRC+18]. Mandatory artifact evaluations [CP15] and artifact databases [BBE+22] further intend to improve the availability of research artifacts.

Still, even work that has been awarded an artifact badge might conceal specific details or limitations [Zil20]. Moreover, artifact evaluations usually fail to fully cover the used methodology. Consequently, the work might still suffer from evaluation inadequacies [vdKAB+18, AQP+22]. Apart from unintentional misconduct, expectations for successful research might be a supporting factor for such improper conduct [LBM10]. Nowadays, the first initiatives started to encourage the reporting of negative results [LAS13, Per22]. Based on this overview, our key takeaway is that even if artifacts are (i) evaluated, (ii) properly documented, (iii) badged, and (iv) open-sourced, they might still not entail the claimed functionality.

Despite the best practices of proposing universally-applicable approaches and preparing well-documented and reusable artifacts, in our interdisciplinary work, we noticed a lack of guidelines that explain the course of action to researchers and practitioners from other domains. In an effort to ease collaborations for future work and to allow for simpler dissemination of new approaches in real-world deployments [MBLT13], we thus derive an abstract methodology for conducting interdisciplinary research. Even though our methodology builds on our experience of interdisciplinary research in the IIoT, we are confident that other areas and domains can equally benefit since the steps of our methodology are not specific to research in the IIoT.

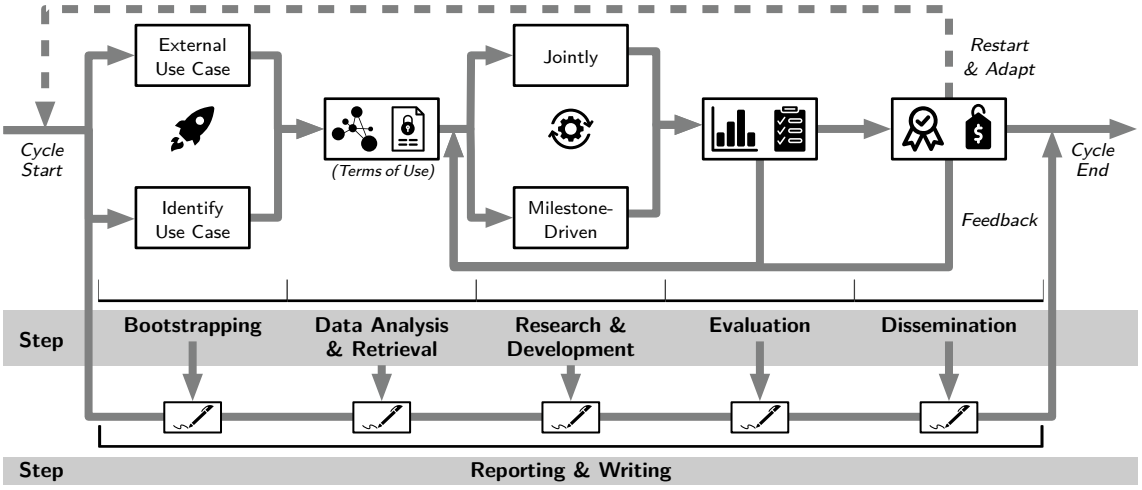



**Figure 6.2** Based on our work, we derived an abstract process cycle consisting of six process steps to formalize the different phases of an interdisciplinary research project at the intersection of cybersecurity research and industrial applications. While the majority of steps, i.e., *Bootstrapping*, *Data Analysis & Retrieval*, *Research & Development*, *Evaluation*, and *Dissemination*, build upon each other, the *Reporting & Writing* step is an accompanying phase.

## 6.2.2 Methodology: A Process Cycle on Research Collaborations

Based on our work in the context of this dissertation and additional research in the context of the IoP, we derived a process cycle that expresses a methodology when conducting interdisciplinary research at the intersection of security research and real-world use cases in industry. Thereby, we intend to support researchers in the organization of their work. In the following, we first provide a high-level overview (Section 6.2.2.1) before detailing the individual steps of our process cycle (Section 6.2.2.2). In Section 6.2.2.3, we conclude our presentation on this excursus with a look at important takeaways and additional lessons learned.

### 6.2.2.1 Design Overview

Based on our experience, we were able to identify patterns that repeatedly occurred as part of our research progress. We derived a (universal) process cycle to guide research in this area. Overall, we separate the cycle into *six* distinct steps, as we illustrate in Figure 6.2. We briefly describe the individual steps in the following.

**Bootstrapping.** First, a suitable use case to work on must be selected. Here, an externally-requested use case can be chosen, or alternatively, a use case can be identified independently. In the latter case, a significant challenge is to make sure that the selected use case is a relevant real-world application, i.e., it is worth pursuing, as well as evaluable and deployable in realistic industry settings.

**Data Analysis & Retrieval.** Second, researchers must make themselves familiar with any available data (sources). This process can also cover the reuse of data and artifacts. Having access to meaningful information is key to developing use case-tailored

solutions. To this end, researchers might even have to accept specific terms of use or sign a non-disclosure agreement to comply with data usage policies, which protect sensitive (real-world) use case data and impairs or even prevents the publication of research artifacts at a later time. Additionally, apart from understanding the semantics, data usually requires pre-processing for subsequent development.

**Research & Development.** Third, a suitable approach for the current use case must be developed. Depending on the type of collaboration, researchers and the use case provider can work jointly, i.e., they have a close feedback loop. Alternatively, the process can be milestone-driven, where progress is only reported periodically.

**Evaluation.** Fourth, researchers must look at the applicability and feasibility of proposed approaches. To this end, they should conduct realistic, real-world-motivated evaluations that convey practical implications and uncover remaining limitations.

**Dissemination.** Fifth, the dissemination of research results, conclusions, lessons learned, methodologies, artifacts, and datasets is an essential step to (i) raise awareness of this area, (ii) encourage additional research, and (iii), most importantly, further utilize the developed approach. In this context, established practices (cf. Section 6.2.1.2), such as public artifacts sharing, come to mind. Especially concerning the *data analysis & retrieval* process step, this phase has a significant impact as publicly-shared (and ideally independently-audited) artifacts can be reused in other contexts. Apart from being a well-acclaimed contribution, sharing artifacts also supports incremental research as the tiresome task of repeating evaluations or re-implementing previous concepts is avoided. At the same time, the overall comparability of results is improved. Thus, it eases the use of standardized methodologies.

**Reporting & Writing.** Finally, we identified a step that accompanies all previously-presented phases. On the one hand, researchers have to report on their recent steps to drive the discourse with their collaborators. On the other hand, they have to keep the (public) dissemination of the approach and the results in mind to receive external feedback (e.g., through peer review) and to fuel the scientific discourse. Most notably, in this step, both researchers and collaborators are involved alike.

Once this process cycle has been completed, research can tackle another use case while also incorporating the newly-generated knowledge and experience. As such, every past use case also influences future challenges and potentially contributes to their resolutions. Next, we detail the individual steps of our process cycle.

#### 6.2.2.2 Revisiting the Process Cycle

Following this abstract overview of our proposed methodology, we now elaborate on the individual steps as well as their underlying aspects. During this presentation, we repeatedly refer to our experience. In accordance with Figure 6.2, we further **highlight** and discuss major alternative actions of individual process steps. For a detailed case study that explores the process cycle in the context of a single contribution (**ACROSS-Matching**), we refer to our previous paper [PBD+21].

### Bootstrapping

Tackling a relevant real-world problem starts with finding such a problem before envisioning a solution for improvements. Although both steps usually require domain knowledge, researchers can choose to actively collaborate with practitioners (*External Use Case*) or bootstrap the work on their own (*Identify Use Case*). Here, the main challenges are that researchers and practitioners might not share the same visions, and they might not immediately understand each other on a technical level. Thus, both the language and the conceptual research ideas must be carefully bridged.

**External Use Case** Given our background, we were in talks with practitioners from the domain of injection molding to look at their data security-related research problems. There, we jointly identified the research challenges of **ACROSS-Comparison** and **ACROSS-Matching**, i.e., a lack of sufficient yet real-world feasible data security for industrial use. While today's security possibilities were mostly unclear for the practitioners, we, as researchers, were initially unable to estimate their needs accurately. A particular challenge in finding a suitable solution is that conservative companies may lack the vision to advance the state of the art using external data.

**Identify Use Case** Regardless, research can also succeed without an initial exchange with a practitioner. For example, **ALONG-Pipeline** is based on our (independent) analysis of related work. Thus, instead of directly talking with practitioners, we only ingested their views indirectly. However, as we already experienced during previous research efforts [HHS+18, SWGW20], this step is cumbersome and might result in identifying research gaps that do not fit real-world industry needs. Hence, all available means must be taken to verify the open problem. Therefore, for **ALONG-Pipeline**, we conducted expert interviews on a small scale after having postulated the research gap internally.

Finally, as part of this step, a thorough analysis of related work is needed. This analysis is not limited to suitable existing solutions, but more importantly, researchers should also try to identify similar (real-world) use cases to consider and improve the universality of their solutions, as expected by academia [PBD+21].

### Data Analysis & Retrieval

After exploring the problem space, any required data must be discovered, shared, and analyzed for later use, as we detail in the following. Depending on the collaboration partners, this step might also include signing usage agreements with them.

**Discovery.** First, researchers have to discover which data is available and then identify which parts of it are relevant. Afterward, they must check whether this information is also accessible for research use and under which terms. For example, further use and access to past benchmarking data (**ACROSS-Comparison**) were limited to the practitioner due to its sensitivity. Again, researchers should be open and try to gather as much information as possible. Practitioners might overlook seemingly-insignificant data as they could be biased by their day-to-day tasks.

Researchers can ideally complement (sensitive) data from practitioners with public datasets to foster reproducibility by and comparability to other work. However, such

datasets also introduce their own caveats. For instance, in addition to the limited number of available datasets for cybersecurity research of industrial systems, they are usually either not a perfect fit for the use case at hand or only contain poor documentation of the underlying setup [CDT21]. Thus, as incorporating "correct" and meaningful data is crucial, we urge researchers to carefully review their selection process and to expect conceptual flaws due to the origins of (real-world) data.

**Sharing.** Following the discovery, when data is not publicly available, the information must be transferred to the researchers. Most notably, the main issue is usually not whether data exists but whether it could be shared for external research purposes, i.e., confidentiality concerns challenge its use. Thus, any questions concerning the terms of use must be (formally) resolved (cf. **Terms of Use** below).

Our work on securing collaborations is especially challenged by the fact that receiving real-world data from *different* companies is mostly impossible as they fear data leaks and a loss of their competitive advantage (cf. Section 1.1). Thus, during development and evaluation, we usually have to split existing data or "reuse" the same dataset for all involved parties in our setting. A related aspect that is especially prevalent in industrial settings is the challenge of real-world implications, i.e., does our data use have any implications on productive systems. For example, as part of Internet measurements [DLF+20, DLP+22], we directly interacted with real-world deployments. Thus, we had to ensure that our research had no unexpected and undesired real-world implications (mainly concerning safety and security).

**Analysis.** After having access to the data, understanding its semantics and structure is the next challenge. For **ACROSS-Matching**, we initially had to work through the data without any digital or paper-based documentation, i.e., we had to collaborate closely with the corresponding practitioner to correctly interpret the data. Apart from identifying gaps in the available information, the use case data must be translated into a proper form, where unnecessary information is removed. Given that practitioners might not be familiar with the used data sources either, this process can become quite time-consuming until a correct pre-processing has been applied. In addition to accurate insights, this step also allows researchers to transform the information into representations in open and standardized formats.

Naturally, a correct understanding of the available information is key to avoiding subsequent errors. However, achieving this state is difficult as often, no or little documentation is available. Industry data frequently originates from proprietary systems that impose additional obstacles to this step. Already the first glance might be misleading, which invalidates all subsequent steps of the research.

**Terms of Use** Industrial companies are notoriously conservative due to their competitive standing: Data is both valuable (also as part of global collaborations) and sensitive at the same time. Therefore, we noticed that discussions regarding potential improvements fueled by security research are usually very enthusiastic. However, they are frequently reserved when being asked to share data or open-source research artifacts, requiring researchers to accept any terms of use and/or to sign non-disclosure agreements. Different best practices in companies and academia can further delay this already slow process. For our first evaluations of **ALONG-Pipeline**,

we did not rely on company data as we initially generated artificial datasets on real-world supply chains for the development. Thus, we did not encounter any delays from the obstacles of signing a non-disclosure agreement. However, for other settings, coming up with realistic or even usable artificial datasets might not be an option. Thus, getting access to otherwise confidential (protected) use case data unlocks otherwise unavailable research challenges and improves the real-world focus.

#### Research & Development

Once the needed use case data is prepared, work on the appropriate solution can start. While this process does not differ research-wise, we observed that the way of interaction with the use case partner can affect the progress. In particular, we distinguish a joint development approach from a more milestone-driven paradigm.

**Jointly** Collaboratively conducting an agile process is extremely helpful in correcting any newly-occurring (or remaining) misconceptions early on. Furthermore, this approach allows researchers to demonstrate any increments while also raising awareness of the associated technical challenges. Thus, it fosters the ongoing discourse. Despite a thorough data analysis & retrieval phase, during our work on **ACROSS-Matching**, we still discovered a misconception concerning the use case data. Fortunately, our close collaboration quickly allowed us to adjust the development accordingly.

**Milestone-Driven** Alternatively, the reporting can be limited to specific milestones. Here, the benefit is that the practitioner can focus on the use case and is not repeatedly distracted by arbitrary technical details. However, receiving timely and accurate feedback on recent progress is more challenging when pursuing this approach. For our supply chain work (**ALONG-Pipeline**), we finished a first prototype without having access to any use case data. Thus, we had no feedback regarding the real-world applicability or even correctness. Apart from the risk of solving the wrong real-world "problem", we noticed that obtaining suitable evaluation data can become extremely tiresome, e.g., if a non-disclosure agreement needs to be signed. Besides, a lack of real-world data complicates the publication of the work.

Overall, we want to highlight that having a feedback loop in place is very beneficial, given that any project on secure industrial collaborations puts a strong emphasis on real-world applicability. We valued the iterative development approach while designing our parameter exchange (**ACROSS-Matching**) because it allowed us to correct misconceptions early on. Hence, ensuring that this key goal is met also helps to confirm the correctness of the solution without additional overhead. Finally, the practitioners will feel more integrated into the research project, which reduces the risks of indifference or dissatisfaction. In the past, we heard reports from practitioners in similar interdisciplinary projects that they feel to just serve as data sources, i.e., they felt underrated concerning their contribution to the research progress.

Another major challenge that we repeatedly came across as part of our work concerns the scalability requirements of the approach in question. While the evaluation itself follows as the next step in our process cycle, realistic constraints are essential upfront to come up with a fitting solution. We frequently noticed that the exact future needs

are still unclear as the overall future development concerning data-sharing benefits and data-security demands for this comparably-novel research intersection are mostly uncertain. Paired with the reservation toward change in these usually-conservative environments, correctly inferring the scalability needs is difficult.

### Evaluation

At this point, we want to explicitly highlight the need to check for real-world applicability during evaluations. Overall, every evaluation should indicate whether the developed prototype is suitable to tackle the targeted real-world problem and which consequences the results entail. If needed, the developed approach must be revised thoroughly according to new findings during the evaluation.

As the main goal of the evaluation is to show that the developed prototype meets real-world requirements and to avoid any inaccurate conclusions, we recommend relying on real-world use case data at all times. For **ALONG-Pipeline** and **ACROSS-Matching**, we were even able to discover additional use cases and acquire their data while evaluating the original use case, which forced us to also thoroughly analyze and interpret this data. While already traditional evaluations in the privacy research area can take significant time, high volumes of real-world data can increase the time beyond that. Rather, however, the need to access or interact with industrial machines (that are used in production) may impact the duration of this phase. Naturally, as in all other steps, specific caution must be exercised concerning possible safety aspects and environmental impact [BDJ+22]. To further improve the impact of publications in terms of security research, we recommend conducting use case-independent evaluations as well, i.e., to generalize the security contributions as much as possible and to demonstrate their overall scalability for applications beyond the specific use case. This aspect supports researchers who are challenged with deriving claims that are universally valid or gathering all-encompassing empirical evidence. For all of our approaches, we generated large, artificial datasets based on real-world use case data to explore the limits of our implementations.

### Dissemination

Especially with the focus on real-world-applicable solutions, steps to disseminate the progress are crucial. Here, we identified different aspects where the interests of researchers and practitioners diverge, e.g., the usability of a prototype developed during research or publishing use case data. Hence, researchers must discuss and agree on these aspects with all stakeholders early on to avoid misconceptions.

**Readiness Level.** Keeping in mind that research is usually only interested in developing proof-of-concept prototypes, the trade-off between their usability and the impact on research needs consideration. Especially with practitioners as partners who strive for real-world deployments of said developments, the expectations should be clarified at the start of the cooperation, i.e., when beginning the process cycle for the first time. We believe that the contribution of convincing practitioners to be open-minded for novel approaches that are enabled by security research is already of intangible value despite a potentially limited product maturity.

**Preparing Artifacts.** Orthogonally, the publication of datasets is a delicate aspect as they can still contain sensitive use case information. A non-disclosure agreement, which covers the use case data of our second evaluation of **ALONG-Pipeline**, prevents us from publishing the corresponding information in any form. In contrast, for **ACROSS-Matching**, we communicated our desire to prepare a public artifact early on. Thus, the practitioners could ensure that the evaluation data did not contain any sensitive data. As removing all critical, potentially insights-leaking features is very challenging, we can understand the reservations of companies and other stakeholders when it comes to their otherwise private data. Relatedly, we had to strip any privacy-sensitive or de-anonymizing user data from compiled Internet measurements' datasets [Data20, Data21] before their open-sourcing to avoid misuse.

**Reusability.** Apart from verifiability, artifacts should also improve reusability. However, preparing an all-encompassing documentation is far from trivial. This process is further challenged if arbitrary use case data should be supported, as individual data sources can vary significantly in syntaxes and semantics. Here, domain-induced misconceptions might challenge attempts to pre-process data correctly.

The research community already looks into ways to improve the status quo and offers programs, such as artifact evaluations, badges, and other awards (cf. Section 6.2.1.2). For example, for **ACROSS-Matching**, we open-sourced our implementation and all use case data [SrcC20] and further received a functional badge [PBL+20]. With a large number of available artifacts, the likelihood of defining (and reusing) a standardized research methodology across use cases, domains, and academia increases. As for all research in general, advances building upon existing approaches can, in the long run, also help to tackle problems that seem unsolvable at the moment.

**Responsible Disclosure.** Nevertheless, when researchers discover alarming information, ethical principles require them to actively and responsibly disclose the relevant findings. As part of Internet-wide measurements of the security configuration of industrial deployments, we reached out to operators whenever possible [DLF+20].

**Bootstrapping Further Research.** All these different aspects of the dissemination can help to ensure progress. Eventually, we are confident that any work in this novel research intersection can encourage additional work, resulting in a larger overall acceptance of this challenging yet practical area. Regardless, with any finished project, researchers can now revisit other use cases and build upon the newly-gathered experience. For example, previous work [RDF+20] sparked the idea for our Internet measurements. With real-world applicable designs and deployments in mind, reporting on negative findings is supportive for the research community as well.

#### Reporting & Writing

Reporting on the research progress and publishing results is as important as conducting the (applied) research. We consider this process step to evolve in parallel to all previously-presented steps as each step provides meaningful input to it.

As part of our collaborations, we noticed various challenges with scientific writing when many stakeholders from different domains are involved, as each stakeholder

is accustomed to their individual best practices or tooling (e.g., the use of LaTeX or versioning systems). These differences require an alignment across stakeholders. However, the challenges also comprise expectations regarding (writing) styles, submission processes, and other organizational matters. At this point, we also want to explicitly raise the need for strict compliance with established approval processes. Thus, all expectations and deadlines should be communicated clearly and early on. For our **ALONG-Finding** paper, we had to manage nine authors from five departments, all contributing their own publication cultures, expectations, and processes.

#### 6.2.2.3 Takeaways and Other Lessons Learned

In addition to the aspects that are attributable to specific steps in our process cycle, we also experienced some additional lessons learned that do not fit into a single step.

**Communication.** As our research focuses on interdisciplinary research topics, an active exchange among all stakeholders is very important. To detail our corresponding experience, we previously described them in the context of our work on **ACROSS-Matching** [PBD+21]. Generally, we noticed that while the first discussions are challenging to master, the situation improves over time as the awareness of the motivation, challenges, and fears is increasingly understood by the collaborators. The goal must be to jointly tackle the issue rather than working individually to resolve implicit assumptions early on that potentially hinder substantial progress.

**Curiosity.** Even though domain experts are usually involved, we noticed that challenging their views, assumptions, and ideas is helpful to deepen the understanding of the topic on the one hand and to revisit the fit of a chosen approach on the other hand. Hence, we recommend questioning everything and not taking anything for granted, as bridging the domains is very challenging and takes significant time.

**Artifact Reuse and Comparability.** As we discussed in Section 6.2.1.2, artifact reuse is an important aspect. Even so, the unavailability of real-world use case data or appropriate models [PMK+24] significantly challenges new research activities. This constraint impairs the comparability with other approaches and software artifacts. We report on these issues in more detail in our previous paper [PBD+21]. In this context, to mitigate and address these challenges, we also elaborate on our vision to transform the *Reporting & Writing* phase into a *Documenting & Writing* phase by integrating concepts from applied RDM.

Based on our activities and collaborations in the IIoT (primarily in the context of contributions of this dissertation), we derived a process cycle that captures any steps that had to be taken and any challenges that had to be tackled. Thereby, we intend to provide a better understanding and formalization of interdisciplinary research efforts in general. By raising the importance of these issues to both researchers and practitioners, we hope to contribute to additional successful research activities, ideas, and visions in the future. Focusing on secure collaborations again, we look forward to novel, more sophisticated, and complex applications not only within the information security dimension but also in the other research dimensions that are crucial to widely evolving the industrial landscape and collaborations in industry.

# 7

# Conclusion

Various trends contributed to the increasingly-interconnected and networked industrial landscape, commonly referred to as the Industrial Internet of Things (IIoT). As a result, companies can (locally) access, process, and analyze vast amounts of information and knowledge to improve their processes and operations. Various initiatives (cf. Section 1.1) and research projects (including the "Internet of Production"; cf. Section 2.1.1) expect that significant improvements can be unlocked by exchanging information globally: To pursue this vision, companies need to exchange (sensitive) information with other stakeholders. Otherwise, they cannot extensively and optimally exploit the available knowledge that is currently encapsulated in local information silos. However, related work has yet to demonstrate that stakeholders can rely on (technical) confidentiality guarantees while sharing sensitive information. To address this gap, in this dissertation, we conceptualized the secure realization of these information flows as secure collaborations. In particular, we thoroughly assessed whether we are able to secure collaborations in real-world settings.

With our four contributions (Chapters 4 and 5), we have demonstrated that we are indeed able to reliably secure collaborations along and across supply chains, both with known and unknown collaborators, using well-established building blocks from private computing. By evaluating several real-world use cases from the domain of production technology, we have further shown that our proposed designs scale to industry-sized applications. Consequently, from the information-security perspective, we have all tools (building blocks) at hand to substantially evolve the industrial landscape by globally establishing secure collaborations in the IIoT. Thereby, we support very diverse goals, from cost reductions over sustainable improvements to reliably-attested fair trade for suppliers, manufacturers, and consumers alike.

In the remainder of this dissertation, we first briefly outline the main takeaways of our work and contributions in Section 7.1. Afterward, in Section 7.2, we highlight the next steps for future work before concluding with some final remarks in Section 7.3.

## 7.1 Takeaways for Secure Collaborations in the IIoT

In this dissertation, we have initially studied the industrial landscape, its entities, and conceivable information flows in the IIoT (cf. Section 2.1.2). Based on this analysis, we have identified collaborations along and across supply chains. We then researched how to reliably secure them using building blocks from private computing. Specifically, we make four primary contributions in this dissertation (cf. Figure 1.4): ① *A Processing Pipeline for Reliable Information* (Chapter 4.1), ② *Finding New Suppliers with Privacy-Preserving Purchase Inquiries* (Chapter 4.2), ③ *Privacy-Preserving Company Benchmarking* (Chapter 5.1), and ④ *Privacy-Preserving Parameter Exchange* (Chapter 5.2). We refer to our presentation in Section 6.1.2 for a corresponding summary of our contributions and focus on the bigger picture now.

Based on this dissertation's findings, our conducted research, and the presented contributions, we conclude that approaches to reliably secure collaborations along supply chains are mature and ready for deployment in today's industrial landscape. Hence, they can immediately support companies in the IIoT. While this form of collaboration is more traditional because the corresponding information sharing aligns itself in a way with the flows of physical goods, we also demonstrated the feasibility of collaborations across supply chains. In particular, we have proposed novel designs that feature technical confidentiality guarantees to realize them convincingly and securely. Thereby, we underline the practicality from an information-security perspective and further prove that the potential of globally sharing information through secure collaborations is well within reach. By relying on well-established building blocks from private computing, we are confident that reliably-secured collaborations can even convince conservative stakeholders because their refusal would keep them from significant benefits and advances.

Overall, our contributions (and this dissertation) help to fuel the (ongoing) evolution of the industrial landscape. We complement related work and large-scale initiatives (e.g., the IDS or GAIA-X) that largely rely on organizational trust guarantees with concepts for secure collaborations that build on technical guarantees. Thereby, we account for the strict confidentiality requirements by stakeholders in the IIoT. Moreover, we provide solid insights into challenges, requirements, and best practices when realizing secure collaborations in the IIoT, both along and across supply chains. In addition to the direct implications of our contributions, the realization of novel collaborations (which we advance with our findings) also promises to spark ideas on how to evolve building blocks from private computing. Eventually, the mutually-reinforcing evolutionary process of enabling building blocks and secured use cases will lead to currently inconceivable applications and use cases.

Altogether, with this dissertation, we have not only shown that realizing secure collaborations in the IIoT is feasible and practical, i.e., globally tapping into today's isolated information silos is possible, but we have also proven that the interdisciplinary development of sophisticated yet appropriate designs for use case-driven secure collaborations can succeed. In particular, the exchange among computer scientists, supply chain experts, and engineers stimulates benefits for all disciplines.

## 7.2 Future Work

Our research in this dissertation is embedded in the IIoT, which introduces various interesting challenges and research directions. Most of these aspects significantly exceed the scope of this work and go beyond our primary research question that focused on the feasibility of secure industrial collaborations in real-world settings (cf. Section 1.2.2). In the following, we briefly highlight the most relevant directions that are in the scope of evolving and possibly transforming secure collaborations.

First, as we have pointed out in Section 6.1.3, the technologies of federated learning [LSTS20] and process mining [van16] have significant potential to extensively transform the state of the art of secure collaborations that we have established in this dissertation. In particular, they promise to link, integrate, and combine knowledge in the industrial landscape more profoundly beyond the scope of this dissertation. As such, their application could also facilitate increased automation concerning the establishment of collaborations, the adaption of processes, and the dissemination of knowledge. Likewise, we look forward to advances related to verifiable computing [DSB17] because they could contribute to making additional use cases accessible for secure collaborations. However, corresponding advances naturally depend on the acceptance of secure collaborations by stakeholders in the IIoT. Moreover, more invasive collaborations will only succeed if sufficient technical guarantees back them.

Second, we have primarily sourced well-established technical building blocks to realize our designs because we studied secure collaborations in light of their feasibility for immediate deployments. However, with this strategic focus, we excluded uncertain future developments in the information security dimension. Specifically, advances in the area of quantum computing will have significant implications on security mechanisms, concepts, and building blocks as we use, configure, and deploy them today [Mos18]. Consequently, in the long run, future work should reiterate how to securely realize industrial collaborations in light of this threat.

Finally, we have to point out that we mainly focused on secure collaborations from the information security dimension. Hence, future work needs to cover all relevant dimensions, i.e., economic, legal, operational security, and interoperability (cf. Figure 1.2). Advances in these dimensions will equally contribute to the success of secure collaborations. Especially when moving toward practical deployments and real-world use, collaboration-associated costs, including their (software) development, standardization, setup, operation, and maintenance, need to be considered and reasonably distributed among the involved (and benefiting) stakeholders in the industrial landscape. This aspect should not be underestimated because energy consumption and acquisition cost of hardware that reliably stores, processes, and communicates vast amounts of information can be considerable factors. Accordingly, together with stakeholders in the IIoT, we need to simultaneously push and support corresponding research activities to facilitate the wide dissemination of collaborations in the evolving industrial landscape. We are confident that such efforts will support advances toward a more sustainable IIoT [BPR+23], while also contributing to fulfilling the United Nation's SDGs.

## 7.3 Concluding Remarks

In this dissertation, we have extensively studied whether we can realize secure collaborations in the IIoT using well-established building blocks while providing technical guarantees. As we have shown for various types of collaborations (e.g., along and across supply chains), we have the concepts and building blocks available to do so, and corresponding designs can even scale to applications in the industrial landscape. Altogether, with our work and findings, we complement large-scale initiatives that approach information sharing in the IIoT from a conceptual level.

Our interdisciplinary research already allowed us to present the outlined findings of this dissertation in various well-acclaimed and prestigious international research communications, where we came across significant interest for our contributions as well as for the overall research direction of secure collaborations. In this context, our most important lesson learned is that secure collaborations in the IIoT (just like the research covering them) can only succeed through interdisciplinary advances and communication. Therefore, we eagerly await advances in the other essential dimensions that prevent the widespread deployment of secure collaborations (as we have introduced them in this dissertation) in the industrial landscape so far.

We look forward to experiencing whether the contributions, findings, and conclusions in this dissertation will stand the test of time once the industrial landscape has evolved extensively. We further believe that transferring our findings to other emerging areas, such as smart grids or digital health, is a reasonable next step since the combination of confidentiality needs and unrealized information sharing is not unique to the IIoT. Over time, synergies might even emerge following the transfer of building blocks, concepts, and methodologies. Hence, our research could also contribute to the evolution of other areas along the way. Jointly with our (research) partners, we are convinced that secure (industrial) collaboration will greatly impact companies, consumers, and society alike in the future. Even if our contributions are superseded at some point, we are confident that our conceptualized research methodology will remain relevant as interdisciplinary cooperation is key to repeatedly advancing the state of the art. Thank you for following our presentation, processing the information in this dissertation, and potentially disseminating it across disciplines and around the world.

The End.

*. . . and the beginning of follow-up research.*

# Glossary

**advanced encryption standard** AES is a secure and performant cipher block-based symmetric encryption method.

**ALICE** Initiative (alliance) for logistics innovation through collaboration in Europe.

**coopetition** Competing businesses cooperate for their individual benefits.

**digitalization** Applying digital technologies to transform (business) processes.

**digitization** Converting analog information into a digital representation.

**finable, accessible, interoperable, reusable** The FAIR principles mandate specific properties for data to optimize its reuse.

**GAIA-X** Initiative to establish a federated data infrastructure in Europe.

**International Data Spaces** Initiative to establish secure cross-domain data spaces for different industries while considering data sovereignty.

**Internet of Production** The only government-funded research cluster at the intersection of production technology and computer science in Germany.

**non-disclosure agreement** A legally-binding contract that ensures confidentiality of sensitive information among 2+ parties.

**Rivest–Shamir–Adleman** RSA is one of the oldest public-key cryptosystems.

**syndicated procurement** Grouping the orders from multiple buyers with the intention to obtain better offers when compared to individual procurement.

# List of Abbreviations and Acronyms

| | |
|---|---|
| **ABE** | attribute-based encryption |
| **API** | application programming interface |
| **CPS** | cyber-physical system |
| **DAG** | directed acyclic graph |
| **E2E** | end-to-end |
| **EPCIS** | electronic product code information services |
| **ESG** | environmental, social, and corporate governance |
| **FAIR** | finable, accessible, interoperable, reusable |
| **FHE** | fully homomorphic encryption |
| **GDPR** | General Data Protection Regulation |
| **HE** | homomorphic encryption |
| **HTTP** | hypertext transfer protocol |
| **IDS** | International Data Spaces |
| **IIoT** | Industrial Internet of Things |
| **IoP** | Internet of Production |
| **IoT** | Internet of Things |
| **KPI** | key performance indicator |
| **LMAS** | line-less mobile assembly system |
| **MaaS** | Manufacturing-as-a-Service |
| **MMIO** | memory-mapped input/output |
| **NDA** | non-disclosure agreement |
| **OPE** | order-preserving encryption |
| **ORE** | order-revealing encryption |
| **OT** | oblivious transfer |
| **PHE** | partially homomorphic encryption |
| **PPC** | production planing and control |
| **PSI** | private set intersection |
| **RDF** | resource description framework |
| **RDM** | research data management |
| **SCMaaS** | Supply-Chain-Management-as-a-Service |
| **SDGs** | Sustainable Development Goals |
| **SWHE** | somewhat homomorphic encryption |
| **TEE** | trusted execution environment |
| **TLS** | transport layer security |
| **URI** | uniform resource identifier |
| **URL** | uniform resource locator |

# Bibliography


[5st12] 5stardata.info. 5-star Open Data. https://5stardata.info/, 2012.

[AA07] Mohsen Attaran and Sharmin Attaran. Collaborative supply chain management: the most promising practice for building efficient and sustainable supply chains. *Business Process Management Journal*, 13(3):390–404, 2007.

[AAUC18] Abbas Acar, Hidayet Aksu, A. Selcuk Uluagac, and Mauro Conti. A Survey on Homomorphic Encryption Schemes: Theory and Implementation. *ACM Computing Surveys*, 51(4):1–35, 2018.

[AB20] Vito Ryan Aprilio and Bettine Bergmans. Implementation of Blockchain for Increasing Traceability at VehGro Supply Chain. *Diponegoro Journal of Accounting*, 9(4), 2020.

[ABL+04] Mikhail Atallah, Marina Bykova, Jiangtao Li, Keith Frikken, and Mercan Topkara. Private Collaborative Forecasting and Benchmarking. In *Proceedings of the 2004 ACM Workshop on Privacy in the Electronic Society (WPES '04)*, pages 103–114. ACM, 2004.

[ACM20] ACM. Artifact Review and Badging - Current. https://www.acm.org/publications/policies/artifact-review-and-badging-current, 2020.

[ADY+19] Emekcan Aras, Stéphane Delbruel, Fan Yang, Wouter Joosen, and Danny Hughes. A Low-Power Hardware Platform for Smart Environment as a Call for More Flexibility and Re-Usability. In *Proceedings of the 2019 International Conference on Embedded Wireless Systems and Networks (EWSN '19)*, pages 194–205. Junction Publishing, 2019.

[AFGH06] Giuseppe Ateniese, Kevin Fu, Matthew Green, and Susan Hohenberger. Improved Proxy Re-encryption Schemes with Applications to Secure Distributed Storage. *ACM Transactions on Information and System Security*, 9(1):1–30, 2006.

[AFS+22] Louise Axon, Katherine Fletcher, Arianna Schuler Scott, Marcel Stolz, Robert Hannigan, Ali El Kaafarani, Michael Goldsmith, and Sadie Creese. Emerging Cybersecurity Capability Gaps in the Industrial Internet of Things: Overview and Research Agenda. *Digital Threats: Research and Practice*, 3(4), 2022.

[Age01] Evgeniy A. Ageshin. E-procurement at work: A case study. *Production and Inventory Management Journal*, 42(1):48–53, 2001.

[AGM+13] Joseph A. Akinyele, Christina Garman, Ian Miers, Matthew W. Pagano, Michael Rushanan, Matthew Green, and Aviel D. Rubin. Charm: a framework for rapidly prototyping cryptosystems. *Journal of Cryptographic Engineering*, 3(2):111–128, 2013.

[AiF98] AiF e.V. (German Federation of Industrial Research Associations). AiF e.V. https://www.aif.de/english/home.html, 1998.

[AIGARB13] Alejandro Alvarado Iniesta, Jorge L García Alcaraz, and ManuelIván Rodríguez Borbón. Optimization of injection molding process parameters by a hybrid of artificial neural network and artificial bee colony algorithm. *Revista Facultad de Ingeniería Universidad de Antioquia*, (67):43–51, 2013.

[AIM10] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The Internet of Things: A survey. *Computer Networks*, 54(15):2787–2805, 2010.

[AK97] Ross Anderson and Markus Kuhn. Low Cost Attacks on Tamper Resistant Devices. In *Proceedings of the 5th International Workshop on Security Protocols (Security Protocols '97)*, volume 1361, pages 125–136. Springer, 1997.

[AKSX04] Rakesh Agrawal, Jerry Kiernan, Ramakrishnan Srikant, and Yirong Xu. Order Preserving Encryption for Numeric Data. In *Proceedings of the 2004 ACM SIGMOD International Conference on Management of Data (SIGMOD '04)*, pages 563–574. ACM, 2004.

[ALSZ13] Gilad Asharov, Yehuda Lindell, Thomas Schneider, and Michael Zohner. More Efficient Oblivious Transfer and Extensions for Faster Secure Computation. In *Proceedings of the 2013 ACM SIGSAC Conference on Computer and Communications Security (CCS '13)*, pages 535–548. ACM, 2013.

[AM16] Saveen A. Abeyratne and Radmehr Monfared. Blockchain ready manufacturing supply chain using distributed ledger. *International Journal of Research in Engineering and Technology*, 5(9), 2016.

[AN07] Rebecca Angeles and Ravi Nath. Business-to-business e-procurement: success factors and challenges to implementation. *Supply Chain Management*, 12(2):104–115, 2007.

[AN18] Anna Adamik and Michał Nowicki. Preparedness of companies for digital transformation and creating a competitive advantage in the age of Industry 4.0. In *In Proceedings of the 17th International Conference on Business Excellence (ICBE '18)*, volume 12, pages 10–24. De Gruyter, 2018.

[AQP+22] Daniel Arp, Erwin Quiring, Feargus Pendlebury, Alexander Warnecke, Fabio Pierazzi, Christian Wressnegger, Lorenzo Cavallaro, and Konrad Rieck. Dos and Don'ts of Machine Learning in Computer Security. In *Proceedings of the 31st USENIX Security Symposium (SEC '22)*, pages 3971–3988. USENIX Association, 2022.

[ATG+16] Sergei Arnautov, Bohdan Trach, Franz Gregor, Thomas Knauth, Andre Martin, Christian Priebe, Joshua Lind, Divya Muthukumaran, Dan O'keeffe, Mark L. Stillwell, David Goltzsche, David Eyvers, Rüdiger Kapitza, Peter Pietzuch, and Christof Fetzer. SCONE: Secure Linux Containers with Intel SGX. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, pages 689–703. USENIX Association, 2016.

[AW08] Alvaro Arenas and Michael Wilson. Contracts as Trust Substitutes in Collaborative Business. *Computer*, 41(7):80–83, 2008.

[Bad20] Lennart Bader. Privacy and Transparency in Digital Supply Chains. Master's thesis, RWTH Aachen University, 2020.

[Bar04] Mark Barratt. Understanding the meaning of collaboration in the supply chain. *Supply Chain Management*, 9(1):30–42, 2004.

[BBC+11] Stefan Behnel, Robert Bradshaw, Craig Citro, Lisandro Dalcin, Dag Sverre Seljebotn, and Kurt Smith. Cython: The Best of Both Worlds. *Computing in Science & Engineering*, 13(2):31–39, 2011.

[BBE+22] David Balenson, Terry Benzel, Eric Eide, David Emmerich, David Johnson, Jelena Mirkovic, and Laura Tinnel. Toward Findable, Accessible, Interoperable, and Reusable Cybersecurity Artifacts. In *Proceedings of the 15th Workshop on Cyber Security Experimentation and Test (CSET '22)*, pages 65–70. ACM, 2022.

[BBM00] Mihir Bellare, Alexandra Boldyreva, and Silvio Micali. Public-Key Encryption in a Multi-user Setting: Security Proofs and Improvements. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT '00)*, pages 259–274. Springer, 2000.

[BBS19] Kilian Becher, Martin Beck, and Thorsten Strufe. An Enhanced Approach to Cloud-based Privacy-preserving Benchmarking. In *Proceedings of the 2019 International Conference on Networked Systems (NetSys '19)*. IEEE, 2019.

[BCD+14] Adam Back, Matt Corallo, Luke Dashjr, Mark Friedenbach, Gregory Maxwell, Andrew Miller, Andrew Poelstra, Jorge Timón, and Pieter Wuille. Enabling Blockchain Innovations with Pegged Sidechains. Technical report, Blockstream, 2014.

[BDCOP04] Dan Boneh, Giovanni Di Crescenzo, Rafail Ostrovsky, and Giuseppe Persiano. Public Key Encryption with Keyword Search. In *Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT '04)*, volume 3027, pages 506–522. Springer, 2004.

[BDJ+22] Philipp Brauner, Manuela Dalibor, Matthias Jarke, Ike Kunze, István Koren, Gerhard Lakemeyer, Martin Liebenberg, Judith Michael, Jan Pennekamp, Christoph Quix, Bernhard Rumpe, Wil van der Aalst, Klaus Wehrle, Andreas Wortmann, and Martina Ziefle. A Computer Science Perspective on Digital Transformation in Production. *ACM Transactions on Internet of Things*, 3(2), 2022.

[BDRW19] Franco Basso, Sophie D'Amours, Mikael Rönnqvist, and Andrés Weintraub. A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research*, 26(3):775–793, 2019.

[Bea96] Donald Beaver. Correlated Pseudorandomness and the Complexity of Private Computations. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC '96)*, pages 479–488. ACM, 1996.

[BEM22] Simon Birnbach, Simon Eberz, and Ivan Martinovic. Haunted House: Physical Smart Home Event Verification in the Presence of Compromised Sensors. *ACM Transactions on Internet of Things*, 3(3), 2022.

[Ben20] Ayoub Benaissa. PyPSI. `https://github.com/OpenMined/PyPSI`, 2020.

[BFR+18] Ferdinand Brasser, Tommaso Frassetto, Korbinian Riedhammer, Ahmad-Reza Sadeghi, Thomas Schneider, and Christian Weinert. VoiceGuard: Secure and Private Speech Processing. In *Proceedings of the 19th Annual Conference of the International Speech Communication Association (Interspeech '18)*, pages 1303–1307. International Speech Communication Association (ISCA), 2018.

[BFRLG21] Arnaud Braud, Gaël Fromentoux, Benoit Radier, and Olivier Le Grand. The Road to European Digital Sovereignty with Gaia-X and IDSA. *IEEE Network*, 35(2):4–5, 2021.

[BGGN+20] Sebastian R. Bader, Irlan Grangel-Gonzalez, Priyanka Nanjappa, Maria-Esther Vidal, and Maria Maleshkova. A Knowledge Graph for Industry 4.0. In *Proceedings of the 17th International Conference on The Semantic Web (ESWC '20)*, volume 12123, pages 465–480. Springer, 2020.

[BGH+13] Dan Boneh, Craig Gentry, Shai Halevi, Frank Wang, and David J. Wu. Private Database Queries Using Somewhat Homomorphic Encryption. In *Proceedings of the 11th International Conference on Applied Cryptography and Network Security (ACNS '13)*, volume 7954, pages 102–118. Springer, 2013.

[BGV12] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (Leveled) fully homomorphic encryption without bootstrapping. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS '12)*, pages 309–325. ACM, 2012.

[BHCW18] Hugh Boyes, Bil Hallaq, Joe Cunningham, and Tim Watson. The industrial internet of things (IIoT): An analysis framework. *Computers in Industry*, 101:1–12, 2018.

[BHFL20] Andreas Balster, Ole Hansen, Hanno Friedrich, and André Ludwig. An ETA Prediction Model for Intermodal Transport Networks Based on Machine Learning. *Business & Information Systems Engineering*, 62(5):403–416, 2020.

[BHS05] Kirsimarja Blomqvist, Pia Hurmelinna, and Risto Seppänen. Playing the collaboration game right—balancing trust and contracting. *Technovation*, 25(5):497–504, 2005.

[BHSS12] Rainer Bourdon, Andreas Hellmann, Jan-Bernd Schreckenberg, and Ralf Schwegmann. Standardized optimization of process and quality by DOE methods — a short manual for injection molding in practice. *Journal of Plastics Technology*, 8(5):525–549, 2012.

[Bit23] Bitkom. Unternehmen wollen Daten nutzen, aber nicht teilen [Companies want to use data, but they do not want to share any]. `https://www.bitkom.org/Presse/Presseinformation/Datenoekonomie-Unternehmen-nutzen-Daten#_`, 2023 (accessed December 4, 2023).

[BKdL+18] Wolfgang Boos, Christoph Maximilian Bernd Kelzenberg, Johan de Lange, Thilo Konrad Schultes, and Max Busch. Erfolgreich Lieferanten Managen im Werkzeugbau. Technical report, WBA Aachener Werkzeugbau Akademie GmbH, 2018.

[BL06] Tim Berners-Lee. Linked Data - Design Issues. `https://www.w3.org/DesignIssues/LinkedData.html`, 2006.

[Blo70] Burton H. Bloom. Space/Time Trade-Offs in Hash Coding with Allowable Errors. *Communications of the ACM*, 13(7):422–426, 1970.

[BLR+15] Dan Boneh, Kevin Lewi, Mariana Raykova, Amit Sahai, Mark Zhandry, and Joe Zimmerman. Semantically Secure Order-Revealing Encryption: Multi-input Functional Encryption Without Obfuscation. In *Proceedings of the 34th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT '15)*, volume 9057, pages 563–594. Springer, 2015.

[BMBJ19] R. Joseph Bensingh, Rajendra Machavaram, Sadayan Rajendra Boopathy, and Chidambaram Jebaraj. Injection molding process optimization of a bi-aspheric lens using hybrid artificial neural networks (ANNs) and particle swarm optimization (PSO). *Measurement*, 134:359–374, 2019.

[BMS21] Matthias Bodenbenner, Benjamin Montavon, and Robert H. Schmitt. FAIR sensor services - Towards sustainable sensor data management. *Measurement: Sensors*, 18, 2021.

[BMW+21] Armin F. Buckhorst, Benjamin Montavon, Dominik Wolfschläger, Melanie Buchsbaum, Amir Shahidi, Henning Petruck, Ike Kunze, Jan Pennekamp, Christian Brecher, Mathias Hüsing, Burkhard Corves, Verena Nitsch, Klaus Wehrle, and Robert H. Schmitt. Holarchy for Line-less Mobile Assembly Systems Operation in the Context of the Internet of Production. *Procedia CIRP*, 99:448–453, 2021. Proceedings of the 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering (ICME '20).

[BOAA+22] Charles Baah, Douglas Opoku Agyeman, Innocent Senyo Kwasi Acquah, Yaw Agyabeng-Mensah, Ebenezer Afum, Kassimu Issau, Daniel Ofori, and Daniel Faibil. Effect of information sharing in supply chains: understanding the roles of supply chain visibility, agility, collaboration on supply chain performance. *Benchmarking: An International Journal*, 29(2):434–455, 2022.

[Boo21] Wolfgang Boos. Production Turnaround — Turning Data into Sustainability. Technical report, RWTH Aachen University, 2021. White Paper.

[BOS+21] David Berdik, Safa Otoum, Nikolas Schmidt, Dylan Porter, and Yaser Jararweh. A Survey on Blockchain for Information Systems Management and Security. *Information Processing & Management*, 58(1), 2021.

[BOT13] Joshua W. S. Brown, Olga Ohrimenko, and Roberto Tamassia. Haze: Privacy-preserving Real-time Traffic Statistics. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL '13)*, pages 540–543. ACM, 2013.

[BP21] Luis Brandao and Rene Peralta. Privacy-Enhancing Cryptography to Complement Differential Privacy. `https://www.nist.gov/blogs/cybersecurity-insights/privacy-enhancing-cryptography-complement-differential-privacy`, 2021. NIST Differential Privacy Blog Series.

[BPH15] Andrew Baumann, Marcus Peinado, and Galen Hunt. Shielding Applications from an Untrusted Cloud with Haven. *ACM Transactions on Computer Systems*, 33(3), 2015.

[BPM+21] Lennart Bader, Jan Pennekamp, Roman Matzutt, David Hedderich, Markus Kowalski, Volker Lücken, and Klaus Wehrle. Blockchain-Based Privacy Preservation for Supply Chains Supporting Lightweight Multi-Hop Information Accountability. *Information Processing & Management*, 58(3), 2021.

[BPM+23] Matthias Bodenbenner, Jan Pennekamp, Benjamin Montavon, Klaus Wehrle, and Robert H. Schmitt. FAIR Sensor Ecosystem: Long-Term (Re-)Usability of FAIR Sensor Data through Contextualization. In *Proceedings of the 21th IEEE International Conference on Industrial Informatics (INDIN '23)*. IEEE, 2023.

[BPR+23] Sebastian Bernhard, Sebastian Pütz, Calvin Röhl, Ralph Baier, Philipp Brauner, Ester Christou, Hannah Dammers, Roman Flaig, Leon M. Gorißen, Jan-Christoph Heilinger, Christian Hinke, István Koren, Dirk Lüttgens, Michael Millan, Kai Müller, Alexander Schollemann, Luisa Vervier, Thomas Gries, Alexander Mertens, Saskia K. Nagel, Frank T. Piller, Günther Schuh, Martina Ziefle, Verena Nitsch, and Carmen Leicht-Scholten. Sustainability in the Internet of Production: Interdisciplinary Opportunities and Challenges. In *Proceedings of the 2023 IEEE International Symposium on Technology and Society (ISTAS '23)*. IEEE, 2023.

[BPT+23] Lennart Bader, Jan Pennekamp, Emildeon Thevaraj, Maria Spiß, Salil S. Kanhere, and Klaus Wehrle. Reputation Systems for Supply Chains: The Challenge of Achieving Privacy Preservation. In *Proceedings of the 20th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '23)*. Springer, 2023. In Press.

[Bra06] Felix Brandt. How to obtain full privacy in auctions. *International Journal of Information Security*, 5(4):201–216, 2006.

[Bre12] Christian Brecher. *Integrative Production Technology for High-Wage Countries*. Springer, 1st edition, 2012.

[BSKC18] Arati Baliga, I. Subhod, Pandurang Kamat, and Siddhartha Chatterjee. Performance Evaluation of the Quorum Blockchain Platform. arXiv:1809.03421, 2018.

[BSW07] John Bethencourt, Amit Sahai, and Brent Waters. Ciphertext-Policy Attribute-Based Encryption. In *Proceedings of the 2007 IEEE Symposium on Security and Privacy (SP '07)*, pages 321–334. IEEE, 2007.

[BSZ22] Philipp Brauner, Anne Kathrin Schaar, and Martina Ziefle. Interfaces, Interactions, and Industry 4.0: A Framework for the User-Centered Design of Industrial User Interfaces in the Internet of Production. In *Human-Technology Interaction: Shaping the Future of Industrial User Interfaces*, pages 361–388. Springer, 2022.

[BTHN96] Sally Blount, Melissa C. Thomas-Hunt, and Margaret A. Neale. The Price Is Right—Or Is It? A Reference Point Model of Two-Party Price Negotiations. *Organizational Behavior and Human Decision Processes*, 68(1):1–12, 1996.

[Buc20] Erik Buchholz. Privacy-Preserving Exchange of Process Parameters. Master's thesis, RWTH Aachen University, 2020.

[Bun21] Bundestag. Act on Corporate Due Diligence Obligations in Supply Chains. Technical report, Federal Republic of Germany, 2021.

[BV14] Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. *SIAM Journal on Computing*, 43(2):831–871, 2014.

[BV19] Manish Mohan Baral and Amitabh Verma. Cloud-Based Intelligent System for Supply Chain Management: A Future Roadmap for SCM Technologies. In *In Proceedings of 3rd International Conference on Nanoelectronics, Circuits and Communication Systems (NCCS '17)*, volume 511, pages 13–24. Springer, 2019.

[BVC01] Pamela Barnes-Vieyra and Cindy Claycomb. Business-to-business E-commerce: models and managerial decisions. *Business Horizons*, 44(3):13–13, 2001.

[BWK17] Jaewook Byun, Sungpil Woo, and Daeyoung Kim. Efficient and privacy-enhanced object traceability based on unified and linked EPCIS events. *Computers in Industry*, 89:35–49, 2017.

[BWW19] Christian Brecher, Marian Wiesch, and Frederik Wellmann. Productivity Increase – Model-based optimisation of NC-controlled milling processes to reduce machining time and improve process quality. *IFAC-PapersOnLine*, 52(13):1803–1807, 2019.

[CAF13] Ruichuan Chen, Istemi Ekin Akkus, and Paul Francis. SplitX: High-performance Private Analytics. In *Proceedings of the ACM SIGCOMM 2013 Conference on SIGCOMM (SIGCOMM '13)*, pages 315–326. ACM, 2013.

[CAP+13] Corrado Costa, Francesca Antonucci, Federico Pallottino, Jacopo Aguzzi, David Sarriá, and Paolo Menesatti. A Review on Agri-food Supply Chain Traceability by Means of RFID Technology. *Food and Bioprocess Technology*, 6:353–366, 2013.

[Cat10] Daniele Catteddu. Cloud Computing: Benefits, Risks and Recommendations for Information Security. In *Proceedings of the Iberic Web Application Security Conference (IBWAS '10)*, volume 72. Springer, 2010.

[CCX+19] Guoxing Chen, Sanchuan Chen, Yuan Xiao, Yinqian Zhang, Zhiqiang Lin, and Ten H. Lai. SgxPectre Attacks: Stealing Intel Secrets from SGX Enclaves via Speculative Execution. *Proceedings of the 2019 IEEE European Symposium on Security and Privacy (EuroS&P '19)*, pages 142–157, 2019.

[CD16] Konstantinos Christidis and Michael Devetsikiotis. Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, 4:2292–2303, 2016.

[CDE+12] James C. Corbett, Jeffrey Dean, Michael Epstein, Andrew Fikes, Christopher Frost, J.Ĵ. Furman, Sanjay Ghemawat, Andrey Gubarev, Christopher Heiser, Peter Hochschild, Wilson Hsieh, Sebastian Kanthak, Eugene Kogan, Hongyi Li, Alexander Lloyd, Sergey Melnik, David Mwaura, David Nagle, Sean Quinlan, Rajesh Rao, Lindsay Rolig, Yasushi Saito, Michal Szymaniak, Christopher Taylor, Ruth Wang, and Dale Woodford. Spanner: Google's Globally-Distributed Database. In *Proceedings of the 10th USENIX Conference on Operating Systems Design and Implementation (OSDI '12)*, pages 251–260. USENIX Association, 2012.

[CDKS21] Hao Chen, Wei Dai, Miran Kim, and Yongsoo Song. Efficient Homomorphic Conversion Between (Ring) LWE Ciphertexts. In *Proceedings of the 19th International Conference on Applied Cryptography and Network Security (ACNS '21)*, volume 12726, pages 460–479. Springer, 2021.

[CDPS+18] Atanu Chaudhuri, Iskra Dukovska-Popovska, Nachiappan Subramanian, Hing Kai Chan, and Ruibin Bai. Decision-making in cold chain logistics using data analytics: a literature review. *The International Journal of Logistics Management*, 29(3):839–861, 2018.

[CDT21] Mauro Conti, Denis Donadel, and Federico Turrin. A Survey on Industrial Control System Testbeds and Datasets for Security Research. *IEEE Communications Surveys & Tutorials*, 23(4):2248–2294, 2021.

[Cer16] Ceresana. Plastic Injection Market Report. Technical report, Ceresana, 2016.

[CFAF17] Sujit Rokka Chhetri, Sina Faezi, and Mohammad Abdullah Al Faruque. Fix the Leak! An Information Leakage Aware Secured Cyber-Physical Manufacturing System. In *Design, Automation & Test in Europe Conference & Exhibition (DATE '17)*, pages 1408–1413. IEEE, 2017.

[CG17] Qiong Cheng and Chong-Zhi Gao. A cloud aided privacy-preserving profile matching scheme in mobile social networks. In *Proceedings of the 2017 IEEE International Conference on Embedded and Ubiquitous Computing (EUC '17)*, volume 2, pages 195–198. IEEE, 2017.

[CG20] Christian Catalini and Joshua S. Gans. Some Simple Economics of the Blockchain. *Communications of the ACM*, 63(7):80–90, 2020.

[CGGI20] Ilaria Chillotti, Nicolas Gama, Mariya Georgieva, and Malika Izabachène. TFHE: Fast Fully Homomorphic Encryption Over the Torus. *Journal of Cryptology*, 33(1):34–91, 2020.

[CGJ+09] Richard Chow, Philippe Golle, Markus Jakobsson, Elaine Shi, Jessica Staddon, Ryusuke Masuoka, and Jesus Molina. Controlling Data in the Cloud: Outsourcing Computation without Outsourcing Control. In *Proceedings of the 2009 ACM Workshop on Cloud Computing Security (CCSW '09)*, pages 85–90. ACM, 2009.

[CGKS95] Benny Chor, Oded Goldreich, Eyal Kushilevitz, and Madhu Sudan. Private Information Retrieval. In *Proceedings of IEEE 36th Annual Foundations of Computer Science (FOCS '95)*, pages 41–50. IEEE, 1995.

[Cha81] David L. Chaum. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. *Communications of the ACM*, 24(2):84–90, 1981.

[Che16] Xiaofeng Chen. Introduction to Secure Outsourcing Computation. *Synthesis Lectures on Information Security, Privacy, and Trust*, 8(2), 2016.

[CHLR18] Hao Chen, Zhicong Huang, Kim Laine, and Peter Rindal. Labeled PSI from Fully Homomorphic Encryption with Malicious Security. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*, pages 1223–1237. ACM, 2018.

[Cis20] Cisco. Cisco Annual Internet Report (2018–2023) White Paper. White paper, Cisco, 2020.

[CJL+20] Ilaria Chillotti, Marc Joye, Damien Ligier, Jean-Baptiste Orfila, and Samuel Tap. CONCRETE: Concrete Operates oN Ciphertexts Rapidly by Extending TfhE. In *Proceedings of the 8th Workshop on Encrypted Computing & Applied Homomorphic Cryptography (WAHC '20)*, pages 57–63. HomomorphicEncryption.org, 2020.

[CKKS17] Jung Hee Cheon, Andrey Kim, Miran Kim, and Yongsoo Song. Homomorphic Encryption for Arithmetic of Approximate Numbers. In *Proceedings of the 23rd International Conference on the Theory and Application of Cryptology and Information Security (ASIACRYPT '17)*, volume 10624, pages 409–437. Springer, 2017.

[CL15] Adina Crainiceanu and Daniel Lemire. Bloofi: Multidimensional Bloom filters. *Information Systems*, 54:311–324, 2015.

[CLP17] Hao Chen, Kim Laine, and Rachel Player. Simple Encrypted Arithmetic Library - SEAL v2.1. In *Proceedings of the 21st International Conference on Financial Cryptography and Data Security (FC '17)*, volume 10323, pages 3–18. Springer, 2017.

[CLWW16] Nathan Chenette, Kevin Lewi, Stephen A. Weis, and David J. Wu. Practical Order-Revealing Encryption with Limited Leakage. In *Revised Selected Papers of the 23rd International Conference on Fast Software Encryption (FSE '16)*, pages 474–493. Springer, 2016.

[CMP+09] George Chryssolouris, Dimitris Mavrikios, Nikolaos Papakostas, Dimitris Mourtzis, George Michalos, and Konstantinos Georgoulias. Digital manufacturing: History, perspectives, and outlook. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 223(5):451–462, 2009.

[CON20] CONSENSYS. ConsenSys Quorum. `https://consensys.net/quorum/`, 2020.

[Con23] ConsenSys, Inc. Comparing proof of authority consensus protocols. `https://docs.goquorum.consensys.net/concepts/consensus/comparing-poa`, 2019 (accessed April 4, 2023).

[CP15] Christian Collberg and Todd Proebsting. A Catalog of Research Artifacts for Computer Science. `http://www.findresearch.org`, 2015.

[CP22] Bhargavi K. Chauhan and Dhirenbhai B. Patel. A Systematic Review of Blockchain Technology to Find Current Scalability Issues and Solutions. In *Proceedings of 2nd Doctoral Symposium on Computational Intelligence (DoSCI '21)*, volume 1374, pages 15–29. Springer, 2022.

[CPR20] Alice Capecchi, Daniel Probst, and Jean-Louis Reymond. One molecular fingerprint to rule them all: drugs, biomolecules, and the metabolome. *Journal of Cheminformatics*, 12(1), 2020.

[CRFG12] Ruichuan Chen, Alexey Reznichenko, Paul Francis, and Johannes Gehrke. Towards Statistical Queries over Distributed Private User Data. In *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation (NSDI '12)*, pages 169–182. USENIX Association, 2012.

[CRZ20] Leon Yang Chu, Ying Rong, and Huan Zheng. The Strategic Benefit of Request for Proposal/Quotation. *Operations Research*, 70(3):1410–1427, 2020.

[CSI14] CSIRO's Data61. python-paillier. `https://github.com/data61/python-paillier`, 2014.

[CT05] Cheng-Kang Chu and Wen-Guey Tzeng. Efficient k-Out-of-n Oblivious Transfer Schemes with Adaptive and Non-adaptive Queries. In *Proceedings of the 8th International Workshop on Theory and Practice in Public Key Cryptography (PKC '05)*, volume 3386, pages 172–183. Springer, 2005.

[Cyr18] Marek A. Cyran. Blockchain as a Foundation for Sharing Healthcare Data. *Blockchain in Healthcare Today*, 1, 2018.

[CZK+19] Raymond Cheng, Fan Zhang, Jernej Kos, Warren He, Nicholas Hynes, Noah Johnson, Ari Juels, Andrew Miller, and Dawn Song. Ekiden: A Platform for Confidentiality-Preserving, Trustworthy, and Performant Smart Contracts. In *Proceedings of the 2019 IEEE European Symposium on Security and Privacy (EuroS&P '19)*, pages 185–200. IEEE, 2019.

[DAAL+17] Stefano De Angelis, Leonardo Aniello, Federico Lombardi, Andrea Margheri, and Vladimiro Sassone. PBFT vs Proof-of-Authority: Applying the CAP Theorem to Permissioned Blockchain. In *Proceedings of the 2nd Italian Conference on Cyber Security (ITASEC '18)*, volume 2058. CEUR Workshop Proceedings, 2017.

[Data20] Markus Dahlmanns, Johannes Lohmöller, Ina Berenice Fink, Jan Pennekamp, Klaus Wehrle, and Martin Henze. Dataset to "Easing the Conscience with OPC UA: An Internet-Wide Study on Insecure Deployments". Dataset RWTH-2020-09197, RWTH Aachen University, 2020.

[Data21] Markus Dahlmanns, Johannes Lohmöller, Jan Pennekamp, Jörn Bodenhausen, Klaus Wehrle, and Martin Henze. Dataset to "Missed Opportunities: Measuring the Untapped TLS Support in the Industrial Internet of Things". Dataset RWTH-2021-10668, RWTH Aachen University, 2021.

[DBN+01] Morris J. Dworkin, Elaine B. Barker, James R. Nechvatal, James Foti, Lawrence E. Bassham, E. Roback, and James F. Dray Jr. Advanced Encryption Standard (AES). NIST FIPS 197, 2001.

[DCLT10] Emiliano De Cristofaro, Yanbin Lu, and Gene Tsudik. Privacy-preserving Sharing of Sensitive Information. Cryptology ePrint Archive 2010/471, 2010.

[DCT10] Emiliano De Cristofaro and Gene Tsudik. Practical Private Set Intersection Protocols With Linear Complexity. In *Proceedings of the 14th International Conference on Financial Cryptography and Data Security (FC '10)*, volume 6052, pages 143–159. Springer, 2010.

[DDJ+20] Volkan Dedeoglu, Ali Dorri, Raja Jurdak, Regio A. Michelin, Roben C. Lunardi, Salil S. Kanhere, and Avelino F. Zorzo. A Journey in Applying Blockchain for Cyberphysical Systems. In *Proceedings of the 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS '20)*, pages 383–390. IEEE, 2020.

[DDM+19] Markus Dahlmanns, Chris Dax, Roman Matzutt, Jan Pennekamp, Jens Hiller, and Klaus Wehrle. Privacy-Preserving Remote Knowledge System. In *Proceedings of the 2019 IEEE 27th International Conference on Network Protocols (ICNP '19)*. IEEE, 2019.

[Dei05] Anil K. Deisingh. Pharmaceutical counterfeiting. *Analyst*, 130(3):271–279, 2005.

[DGH+21] Christoph Dobraunig, Lorenzo Grassi, Lukas Helminger, Christian Rechberger, Markus Schofnegger, and Roman Walch. Pasta: A Case for Hybrid Homomorphic Encryption. Cryptology ePrint Archive 2021/731, 2021.

[DGP03] Antonio Davila, Mahendra Gupta, and Richard Palmer. Moving Procurement Systems to the Internet: the Adoption and Use of E-Procurement Technology Models. *European Management Journal*, 21(1):11–23, 2003.

[DGVPdPS12] Paolo D'Arco, María Isabel González Vasco, Angel L. Pérez del Pozo, and Claudio Soriente. Size-Hiding in Private Set Intersection: Existential Results and Constructions. In *Proceedings of the 5th International Conference on Cryptology in Africa (AFRICACRYPT '12)*, volume 7374, pages 378–394. Springer, 2012.

[DHL+24] Markus Dahlmanns, Felix Heidenreich, Johannes Lohmöller, Jan Pennekamp, Klaus Wehrle, and Martin Henze. Unconsidered Installations: Discovering IoT Deployments in the IPv6 Internet. In *Proceedings of the 2024 IEEE/IFIP Network Operations and Management Symposium (NOMS '24)*. IEEE, 2024. In Press.

[DHRW16] Yevgeniy Dodis, Shai Halevi, Ron D. Rothblum, and Daniel Wichs. Spooky encryption and its applications. In *Proceedings of the 36th Annual International Cryptology Conference (CRYPTO '16)*, volume 9816, pages 93–122. Springer, 2016.

[DJP+19] Volkan Dedeoglu, Raja Jurdak, Guntur D. Putra, Ali Dorri, and Salil S. Kanhere. A Trust Architecture for Blockchain in IoT. In *Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '19)*, pages 190–199. ACM, 2019.

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Proceedings of the 25th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT '06)*, volume 4004, pages 486–503. Springer, 2006.

[DLF+20] Markus Dahlmanns, Johannes Lohmöller, Ina Berenice Fink, Jan Pennekamp, Klaus Wehrle, and Martin Henze. Easing the Conscience with OPC UA: An Internet-Wide Study on Insecure Deployments. In *Proceedings of the ACM Internet Measurement Conference (IMC '20)*, pages 101–110. ACM, 2020.

[DLP+22] Markus Dahlmanns, Johannes Lohmöller, Jan Pennekamp, Jörn Bodenhausen, Klaus Wehrle, and Martin Henze. Missed Opportunities: Measuring the Untapped TLS Support in the Industrial Internet of Things. In *Proceedings of the 17th ACM ASIA Conference on Computer and Communications Security (ASIACCS '22)*, pages 252–266. ACM, 2022.

[DN05] Satyaki Ghosh Dastidar and Rakesh Nagi. Scheduling injection molding operations with multiple resource constraints and sequence dependent setup times and costs. *Computers & Operations Research*, 32(11):2987–3005, 2005.

[DR14] Cynthia Dwork and Aaron Roth. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[DR18] Tim Dierks and Eric Rescorla. The Transport Layer Security (TLS) Protocol Version 1.2. IETF RFC 5246, 2018.

[DSB17] Denise Demirel, Lucas Schabhüser, and Johannes Buchmann. *Privately and Publicly Verifiable Computing Techniques: A Survey.* Springer, 1st edition, 2017.

[DSD+23] Markus Dahlmanns, Constantin Sander, Robin Decker, Klaus Wehrle, Jan Pennekamp, Anastasiia Belova, Thomas Bergs, Matthias Bodenbenner, Andreas Bührig-Polaczek, Ike Kunze, et al. Secrets Revealed in Container Images: An Internet-wide Study on Occurrence and Impact. In *Proceedings of the 2023 ACM Asia Conference on Computer and Communications Security (ASIACCS '23)*, pages 797–811. ACM, 2023.

[Dum07] Graham Dumpleton. Apache/mod_wsgi. `https://github.com/GrahamDumpleton/mod_wsgi`, 2007.

[ECL07] ECLASS e.V. ECLASS – Standard for Master Data and Semantics for Digitalization. `https://www.eclass.eu/`, 2007.

[EGBH15] Dominik Eckstein, Matthias Goellner, Constantin Blome, and Michael Henke. The performance impact of supply chain agility and supply chain adaptability: the moderating effect of product complexity. *International Journal of Production Research*, 53(10):3028–3046, 2015.

[EGL85] Shimon Even, Oded Goldreich, and Abraham Lempel. A Randomized Protocol for Signing Contracts. *Communications of the ACM*, 28(6):637–647, 1985.

[Eij14] Helder Eijs. PyCryptodome. `https://www.pycryptodome.org/`, 2014.

[ElG84] Taher ElGamal. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. *Proceedings of CRYPTO '84*, 196:10–18, 1984.

[ElG85] Taher ElGamal. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. *IEEE Transactions on Information Theory*, 31(4):469–472, 1985.

[Elg12] Fredrik Elgh. Decision support in the quotation process of engineered-to-order products. *Advanced Engineering Informatics*, 26(1):66–79, 2012.

[EPK14] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (CCS '14)*, pages 1054–1067. ACM, 2014.

[Eth18] Ethereum – GitHub. eth-account. `https://github.com/ethereum/eth-account/`, 2018.

[EW17] Michael Ehret and Jochen Wirtz. Unlocking value from machines: business models and the industrial internet of things. *Journal of Marketing Management*, 33(1-2):111–130, 2017.

[FBB15] Gunnar Friede, Timo Busch, and Alexander Bassen. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4):210–233, 2015.

[FG07] Caroline Fontaine and Fabien Galand. A Survey of Homomorphic Encryption for Nonspecialists. *EURASIP Journal on Information Security*, 2007, 2007.

[FHE21] FHE.org. Fully Homomorphic Encryption. `https://fhe.org/`, 2021.

[FHZ10] Barbara B. Flynn, Baofeng Huo, and Xiande Zhao. The impact of supply chain integration on performance: A contingency and configuration approach. *Journal of Operations Management*, 28(1):58–71, 2010.

[FK21] Hajar Fatorachian and Hadi Kazemi. Impact of Industry 4.0 on supply chain performance. *Production Planning & Control*, 32(1):63–81, 2021.

[FNP20] Dario Fiore, Anca Nitulescu, and David Pointcheval. Boosting Verifiable Computation on Encrypted Data. In *Proceedings of the 23rd IACR International Conference on Practice and Theory of Public-Key Cryptography (PKC '20)*, volume 12111, pages 124–154. Springer, 2020.

[Foo18] Food and Drug Administration. Standardization of Data and Documentation Practices for Product Tracing Guidance for Industry. Fda-2018-d-0688, Food and Drug Administration, 2018.

[For23] Foreverhold Ltd. Diamonds. `https://www.everledger.io/industry-solutions/diamonds/`, 2020 (accessed April 4, 2023).

[FPR04] Joan Feigenbaum, Benny Pinkas, and Raphael Ryger. Secure Computation of Surveys. Workshop on Secure Multiparty Protocols (SMP '04), 2004.

[FQS01] FQS e. V. (Research Community Quality). DGQ Forschung – FQS e.V. – Deutsche Geschellschaft für Qualität. `http://www.fqs.de/`, 2001.

[FR01] Michael Funke and Ralf Ruhwedel. Product Variety and Economic Growth: Empirical Evidence for the OECD Countries. *IMF Staff Papers*, 48(2):225–242, 2001.

[FRPO15] Ray Fells, Helen Rogers, Peter Prowse, and Ursula F. Ott. Unraveling Business Negotiations Using Practitioner Data. *Negotiation and Conflict Management Research*, 8(2):119–136, 2015.

[FS00] Paul Ferguson and Daniel Senie. Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing. IETF RFC 2827, 2000.

[Fuh21] Frederik Fuhrmann. Two-way Privacy For Purchase Inquiries in Industry. Master's thesis, RWTH Aachen University, 2021.

[FYDX21] Shufan Fei, Zheng Yan, Wenxiu Ding, and Haomeng Xie. Security Vulnerabilities of SGX and Countermeasures: A Survey. *ACM Computing Surveys*, 54(6), 2021.

[GABAS22] Tan Gürpinar, Sk. Riad Bin Ashraf, Natalia Broza-Abut, and Dominik Sparer. Blockchain-Based Infrastructure for Product Traceability in the Medical Supply Chain. In *Prospects of Blockchain Technology for Accelerating Scientific Advancement in Healthcare*, pages 119–134. IGI Global, 2022.

[GCF11] Saikat Guha, Bin Cheng, and Paul Francis. Privad: Practical privacy in online advertising. In *Proceedings of the 8th USENIX Symposium on Networked Systems Design and Implementation (NSDI '11)*, pages 169–182. USENIX Association, 2011.

[Gen09] Craig Gentry. Fully Homomorphic Encryption Using Ideal Lattices. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC '09)*, pages 169–178. ACM, 2009.

[Gen17] General Assembly. Resolution adopted by the General Assembly on 6 July 2017 – Work of the Statistical Commission pertaining to the 2030 Agenda for Sustainable Development. Technical Report A/RES/71/313, United Nations, 2017.

[GHW+19] René Glebke, Martin Henze, Klaus Wehrle, Philipp Niemietz, Daniel Trauth, Patrick Mattfeld, and Thomas Bergs. A Case for Integrated Data Processing in Large-Scale Cyber-Physical Systems. In *Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS '19)*, pages 7252–7261. AIS, 2019.

[Gil16] Alasdair Gilchrist. *Industry 4.0: The Industrial Internet of Things*. Springer, 1st edition, 2016.

[GKHD20] Peter Gonczol, Panagiota Katsikouli, Lasse Herskind, and Nicola Dragoni. Blockchain Implementations and Use Cases for Supply Chains-A Survey. *IEEE Access*, 8:11856–11871, 2020.

[GKT19] Lampropoulos Georgios, Siakas Kerstin, and Anastasiadis Theofylaktos. Internet of Things in the Context of Industry 4.0: An Overview. *International Journal of Entrepreneurial Knowledge*, 7(1):4–19, 2019.

[GL14] Simson Garfinkel and Heather Richter Lipford. *Usable Security: History, Themes, and Challenges*. Springer, 1st edition, 2014.

[GLD+18] Valentina Gatteschi, Fabrizio Lamberti, Claudio Demartini, Chiara Pranteda, and Víctor Santamaría. Blockchain and Smart Contracts for Insurance: Is the Technology Mature Enough? *Future Internet*, 10(2), 2018.

[Gle23] Lars Christoph Gleim. *An Approach for Global and Local Data Lifecycle Management with Provenance and Persistent Identifiers*. PhD thesis, RWTH Aachen University, 2023.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, 1984.

[GMW91] Oded Goldreich, Silvio Micali, and Avi Wigderson. Proofs that yield nothing but their validity or all languages in NP have zero-knowledge proof systems. *Journal of the ACM*, 38(3):690–728, 1991.

[GN08] Satashu Goel and Rohit Negi. Guaranteeing Secrecy using Artificial Noise. *IEEE Transactions on Wireless Communications*, 7(6):2180–2189, 2008.

[Gor00] Dale K. Gordon. The Past, Present and Future Direction of Aerospace Quality Standards. *Quality Progress*, 33(6):125, 2000.

[GPL+20] Lars Gleim, Jan Pennekamp, Martin Liebenberg, Melanie Buchsbaum, Philipp Niemietz, Simon Knape, Alexander Epple, Simon Storms, Daniel Trauth, Thomas Bergs, Christian Brecher, Stefan Decker, Gerhard Lakemeyer, and Klaus Wehrle. FactDAG: Formalizing Data Interoperability in an Internet of Production. *IEEE Internet of Things Journal*, 7(4):3243–3253, 2020.

[GPSPD06] Angappa Gunasekaran, Goran D Putnik, Josée St-Pierre, and Sylvain Delisle. An expert diagnosis system for the benchmarking of SMEs' performance. *Benchmarking: An International Journal*, 13(1–2):106–119, 2006.

[GPT+21] Lars Gleim, Jan Pennekamp, Liam Tirpitz, Sascha Welten, Florian Brillowski, and Stefan Decker. FactStack: Interoperable Data Management and Preservation for the Web and Industry 4.0. In *Proceedings of the 19th Symposium for Database Systems for Business, Technology and Web (BTW '21)*, volume P-311, pages 371–395. Gesellschaft für Informatik, 2021.

[Gre23] Andy Greenberg. The Untold Story of NotPetya, the Most Devastating Cyberattack in History. `https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world/`, 2018 (accessed April 4, 2023).

[Gro09] Robert L. Grossman. The Case for Cloud Computing. *IT Professional*, 11(2):23–27, 2009.

[GS116] GS1 AISB. EPC Information Services (EPCIS) Standard. Technical Report Release 1.2, GS1, 2016.

[GS121] GS1 AISBL. EPCIS Standard. Technical Report Release 2.0, Community Review Draft, GS1, 2021.

[GSU+22] Gonzalo Munilla Garrido, Johannes Sedlmeir, Ömer Uludağ, Ilias Soto Alaoui, Andre Luckow, and Florian Matthes. Revealing the landscape of privacy-enhancing technologies in the context of data markets for the IoT: A systematic literature review. *Journal of Network and Computer Applications*, 207:103465, 2022.

[GVC+22] Sandra Geisler, Maria-Esther Vidal, Cinzia Cappiello, Bernadette Farias Lóscio, Avigdor Gal, Matthias Jarke, Maurizio Lenzerini, Paolo Missier, Boris Otto, Elda Paja, et al. Knowledge-Driven Data Ecosystems Toward Data Transparency. *Journal of Data and Information Quality*, 14(1), 2022.

[GZZL18] Huang Gao, Yun Zhang, Xundao Zhou, and Dequn Li. Intelligent Methods for the Process Parameter Determination of Plastic Injection Molding. *Frontiers of Mechanical Engineering*, 13(1):85–95, 2018.

[HAAS23] Abhishek Hazra, Mainak Adhikari, Tarachand Amgoth, and Satish Narayana Srirama. A Comprehensive Survey on Interoperability for IIoT: Taxonomy, Standards, and Future Directions. *ACM Computing Surveys*, 55(1), 2023.

[HAPS13] Bernadette Hyland, Ghislain Atemezing, Michael Pendleton, and Biplav Srivastava. Linked Data Glossary. W3C Working Group Note, 2013.

[HBKL20] Christian Hopmann, Pascal Bibow, Thomas Kosthorst, and Yannik Lockner. Process setup in injection moulding by Human-Machine-Interfaces and AI. In *Proceedings of the 30th International Colloquium Plastics Technology*, pages 55–92. Shaker, 2020.

[HBTD21] Patrik Hummel, Matthias Braun, Max Tretter, and Peter Dabrock. Data Sovereignty: A Review. *Big Data & Society*, 8(1), 2021.

[HE02] Karin Höne and Jan Harm Petrus Eloff. Information security policy — what do international information security standards say? *Computers & Security*, 21(5):402–409, 2002.

[Hen20] Martin Henze. The Quest for Secure and Privacy-preserving Cloud-based Industrial Cooperation. In *Proceedings of the 2020 IEEE Conference on Communications and Network Security (CNS '20)*. IEEE, 2020. Proceedings of the 6th International Workshop on Security and Privacy in the Cloud (SPC '20).

[Hep15] Martin Hepp. The Web of Data for E-Commerce: Schema.org and GoodRelations for Researchers and Practitioners. In *Proceedings of the 15th International Conference on Web Engineering (ICWE '15)*, volume 9114, pages 723–727. Springer, 2015.

[HH18] Christian Hopmann and Julian Heinisch. Injection Molding Setup by Means of Machine Learning Based on Simulation and Experimental Data. In *Proceedings of the 76th SPE Annual Technical Conference and Tradeshow (ANTEC '18)*, pages 269–274. Society of Plastics Engineers, 2018.

[HHF+20] Lars Hvam, Christian Lindschou Hansen, Cipriano Forza, Niels Henrik Mortensen, and Anders Haug. The reduction of product and process complexity based on the quantification of product complexity costs. *International Journal of Production Research*, 58(2):350–366, 2020.

[HHH+17] Martin Henze, Jens Hiller, René Hummen, Roman Matzutt, Klaus Wehrle, and Jan Henrik Ziegeldorf. Network Security and Privacy for Cyber-Physical Systems. In *Security and Privacy in Cyber-Physical Systems: Foundations, Principles and Applications*, pages 25–56. Wiley, 2017.

[HHHB11] Hamed Haddadi, Pan Hui, Tristan Henderson, and Ian Brown. Targeted Advertising on the Handset: Privacy and Security Challenges. In *Pervasive Advertising*, pages 119–137. Springer, 2011.

[HHMB19] Fatma Hentati, Ismail Hadriche, Neila Masmoudi, and Chedly Bradai. Optimization of the injection molding process for the PC/ABS parts by integrating Taguchi approach and CAE simulation. *The International Journal of Advanced Manufacturing Technology*, 104(9–12):4353–4363, 2019.

[HHS+16] Martin Henze, Jens Hiller, Sascha Schmerling, Jan Henrik Ziegeldorf, and Klaus Wehrle. CPPL: Compact Privacy Policy Language. In *Proceedings of the 2016 ACM on Workshop on Privacy in the Electronic Society (WPES '16)*, pages 99–110. ACM, 2016.

[HHS+18] Jens Hiller, Martin Henze, Martin Serror, Eric Wagner, Jan Niklas Richter, and Klaus Wehrle. Secure Low Latency Communication for Constrained Industrial IoT Scenarios. In *Proceedings of the 2018 IEEE 43rd Conference on Local Computer Networks (LCN '18)*, pages 614–622. IEEE, 2018.

[HIFZ17] Martin Henze, Ritsuma Inaba, Ina Berenice Fink, and Jan Henrik Ziegeldorf. Privacy-preserving Comparison of Cloud Exposure Induced by Mobile Apps. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '17)*, pages 543–544. ACM, 2017.

[HJD22] Lida Haghnegahdar, Sameehan S. Joshi, and Narendra B. Dahotre. From IoT-based cloud manufacturing approach to intelligent additive manufacturing: industrial Internet of Things—an overview. *The International Journal of Advanced Manufacturing Technology*, 119:1461–1478, 2022.

[HJM+19] Christian Hopmann, Sabina Jeschke, Tobias Meisen, Thomas Thiele, Hasan Tercan, Martin Liebenberg, Julian Heinisch, and Matthias Theunissen. Combined learning processes for injection moulding based on simulation and experimental data. In *Proceedings of the 33rd Polymer Processing Society Annual Meeting (PPS '17)*, volume 2139, pages 152–156. AIP, 2019.

[HLN+07] Wenbo He, Xue Liu, Hoang Nguyen, Klara Nahrstedt, and Tarek Abdelzaher. PDA: Privacy-Preserving Data Aggregation in Wireless Sensor Networks. In *Proceedings of the 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, pages 2045–2053. IEEE, 2007.

[hom17] homomorphicencryption.org. Homomorphic Encryption Standardization. `https://homomorphicencryption.org/`, 2017.

[Hor01] Laura Horvath. Collaboration: the key to value creation in supply chain management. *Supply Chain Management*, 6(5):205–207, 2001.

[HP01] Sung Ho Ha and Sang Chan Park. Matching Buyers and Suppliers: An Intelligent Dynamic-Exchange Model. *IEEE Intelligent Systems*, 16(4):28–40, 2001.

[HP17] Niels Hackius and Moritz Petersen. Blockchain in Logistics and Supply Chain: Trick or Treat? In *Proceedings of the Hamburg International Conference of Logistics (HICL '17)*, volume 23, pages 3–18. epubli, 2017.

[HPD+19] Jens Hiller, Jan Pennekamp, Markus Dahlmanns, Martin Henze, Andriy Panchenko, and Klaus Wehrle. Tailoring Onion Routing to the Internet of Things: Security and Privacy in Untrusted Environments. In *Proceedings of the 2019 IEEE 27th International Conference on Network Protocols (ICNP '19)*. IEEE, 2019.

[HPH+17] Martin Henze, Jan Pennekamp, David Hellmanns, Erik Mühmer, Jan Henrik Ziegeldorf, Arthur Drichel, and Klaus Wehrle. CloudAnalyzer: Uncovering the Cloud Usage of Mobile Apps. In *Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous '17)*, pages 262–271. ACM, 2017.

[HR10] Martin Hepp and Andreas Radinger. eClassOWL – The Web Ontology for Products and Services. `http://www.heppnetz.de/projects/eclassowl/`, 2010.

[HRS+17] Andreas Haas, Andreas Rossberg, Derek L. Schuff, Ben L. Titzer, Michael Holman, Dan Gohman, Luke Wagner, Alon Zakai, and J. F. Bastien. Bringing the web up to speed with WebAssembly. In *Proceedings of the 38th ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI '17)*, pages 185–200. ACM, 2017.

[HRS18] Ali Hasnain and Dietrich Rebholz-Schuhmann. Assessing FAIR Data Principles Against the 5-Star Open Data Principles. In *Proceedings of the ESWC 2018 Satellite Events on the Semantic Web (ESWC '18)*, volume 11155, pages 469–477. Springer, 2018.

[HS14] Kai Hüschelrath and Heike Schweitzer. *Public and Private Enforcement of Competition Law in Europe*. Springer, 1st edition, 2014.

[HSF+09] Dominik Herrmann, Florian Scheuer, Philipp Feustel, Thomas Nowey, and Hannes Federrath. A Privacy-Preserving Platform for User-Centric Quantitative Benchmarking. In *Proceedings of the 6th International Conference on Trust, Privacy and Security in Digital Business (TrustBus '09)*, pages 32–41. Springer, 2009.

[Hue19] Huelse. SEAL-Python. `https://github.com/Huelse/SEAL-Python/`, 2019.

[HX06] Yu Hua and Bin Xiao. A Multi-attribute Data Structure with Parallel Bloom Filters for Network Services. In *Proceedings of the 13th International Conference on High Performance Computing (HiPC '06)*, volume 4297, pages 277–288. Springer, 2006.

[IBM23] IBM Research. Changing the way the world works: IBM Research's "5 in 5". `https://research.ibm.com/blog/ibm-research-5-in-5-2018`, 2019 (accessed April 4, 2023).

[IKNP03] Yuval Ishai, Joe Kilian, Kobbi Nissim, and Erez Petrank. Extending Oblivious Transfers Efficiently. In *Proceedings of the 23rd Annual International Cryptology Conference (CRYPTO '03)*, volume 2729, pages 145–161. Springer, 2003.

[Ind08] Roman Inderst. Single sourcing versus multiple sourcing. *The RAND Journal of Economics*, 39(1):199–213, 2008.

[Int19] International Data Spaces Association. IDS Reference Architecture Model. Technical report, 2019.

[Int23] Intel Corporation. Intel SGX Product Offerings. `https://www.intel.com/content/www/us/en/architecture-and-technology/sgx-product-offerings.html`, 2022 (accessed April 4, 2023).

[ISV19] Sana Imtiaz, Ramin Sadre, and Vladimir Vlassov. On the Case of Privacy in the IoT Ecosystem: A Survey. In *Proceedings of the 12th IEEE International Conference on Internet of Things (iThings '19)*, pages 1015–1024. IEEE, 2019.

[Jar20] Matthias Jarke. Data Sovereignty and the Internet of Production. In *Proceedings of the 32nd International Conference on Advanced Information Systems Engineering (CAiSE '20)*, volume 12127, pages 549–558. Springer, 2020.

[JBSR17] Sabina Jeschke, Christian Brecher, Houbing Song, and Danda B. Rawat. *Industrial Internet of Things: Cybermanufacturing Systems*. Springer, 1st edition, 2017.

[Jes21] Fabian Thorsten Jess. Enhancing Supply Chain Management with Trustworthy and Reliable Sensor Data. Bachelor's thesis, RWTH Aachen University, 2021.

[JHC21] Matthew D. Jones, Scott Hutcheson, and Jorge D. Camba. Past, present, and future barriers to digital transformation in manufacturing: A review. *Journal of Manufacturing Systems*, 60:936–948, 2021.

[JLM88] Van Jacobson, Craig Leres, and Steven McCanne. TCPDUMP/LIBPCAP public repository. `https://www.tcpdump.org/`, 1988.

[JMB11] Sonia Jahid, Prateek Mittal, and Nikita Borisov. EASiER: Encryption-based access control in social networks with efficient revocation. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security (ASIACCS' 11)*, pages 411–415. ACM, 2011.

[JPM16] JP Morgan. Quorum. `https://www.goquorum.com/`, 2016.

[JSB+18] Matthias Jarke, Günther Schuh, Christian Brecher, Matthias Brockmann, and Jan-Philipp Prote. Digital Shadows in the Internet of Production. *ERCIM News*, 115:26–28, 2018.

[KB06] Joakim Kalvenes and Amit Basu. Design of Robust Business-to-Business Electronic Marketplaces with Guaranteed Privacy. *Management Science*, 52(11):1721–1736, 2006.

[KBL18] Djamel Eddine Kouicem, Abdelmadjid Bouabdallah, and Hicham Lakhlef. Internet of things security: A top-down survey. *Computer Networks*, 141:199–221, 2018.

[Ker07] Florian Kerschbaum. Building a Privacy-Preserving Benchmarking Enterprise System. In *Proceedings of the 11th IEEE International Enterprise Distributed Object Computing Conference (EDOC '07)*. IEEE, 2007.

[Ker08] Florian Kerschbaum. Practical Privacy-Preserving Benchmarking. In *Proceedings of The IFIP TC-11 23rd International Information Security Conference (SEC '08)*, pages 17–31. Springer, 2008.

[Ker10] Florian Kerschbaum. *A Privacy-Preserving Benchmarking Platform*. PhD thesis, Karlsruhe Institute of Technology, 2010.

[Ker11] Florian Kerschbaum. Secure and Sustainable Benchmarking in Clouds. *Business & Information Systems Engineering*, 3(3):135–143, 2011.

[KHD17] Kari Korpela, Jukka Hallikas, and Tomi Dahlberg. Digital Supply Chain Transformation toward Blockchain Integration. In *Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS '17)*, pages 4182–4191. AIS, 2017.

[KHH+18] Seongmin Kim, Juhyeng Han, Jaehyeong Ha, Taesoo Kim, and Dongsu Han. SGX-Tor: A Secure and Practical Tor Anonymity Network With SGX Enclaves. *IEEE/ACM Transactions on Networking*, 26(5):2174–2187, 2018.

[KIL09] Musa R. Kamal, Avram I. Isayev, and Shih-Jung Liu. *Injection Molding: Technology and Fundamentals*. Hanser, 2009.

[KJJ99] Paul Kocher, Joshua Jaffe, and Benjamin Jun. Differential Power Analysis. In *Proceedings of the 19th Annual International Cryptology Conference (CRYPTO '99)*, volume 1666, pages 388–397. Springer, 1999.

[KJPE23] István Koren, Matthias Jarke, Frank Piller, and Hoda ElMaraghy. Sustainability and Resilience in Alliance-Driven Manufacturing Ecosystems: A Strategic Conceptual Modeling Perspective. In *Proceedings of the 56th Hawaii International Conference on System Sciences (HICSS '23)*, pages 3653–3662. AIS, 2023.

[KK09] Fritz Klocke and Aaron Kuchle. *Manufacturing Processes 1*. Springer, 2009.

[KKKM13] Abdul Nasir Khan, ML. Mat Kiah, Samee U. Khan, and Sajjad A. Madani. Towards secure mobile cloud computing: A survey. *Future Generation Computer Systems*, 29(5):1278–1299, 2013.

[KKRT16] Vladimir Kolesnikov, Ranjit Kumaresan, Mike Rosulek, and Ni Trieu. Efficient Batched Oblivious PRF with Applications to Private Set Intersection. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security (CCS '16)*, pages 818–829. ACM, 2016.

[KKWF16] Achim Kampker, Kai Kreiskoether, Johannes Wagner, and Sarah Fluchs. Mobile Assembly of Electric Vehicles: Decentralized, Low-Invest and Flexible. *International Journal of Mechanical and Mechatronics Engineering*, 10(12):1976–1982, 2016.

[KL10] Seny Kamara and Kristin Lauter. Cryptographic Cloud Storage. In *Proceedings of the 11th International Conference on Financial Cryptography and Data Security (FC '10)*, volume 6054, pages 136–149. Springer, 2010.

[KL15] Miran Kim and Kristin Lauter. Private genome analysis through homomorphic encryption. *BMC Medical Informatics and Decision Making*, 15 (Suppl 5), 2015.

[KL18a] Merve Can Kus Khalilov and Albert Levi. A Survey on Anonymity and Privacy in Bitcoin-Like Digital Cash Systems. *IEEE Communications Surveys & Tutorials*, 20(3):2543–2585, 2018.

[KL18b] Henry M. Kim and Marek Laskowski. Toward an ontology-driven blockchain design for supply-chain provenance. *Intelligent Systems in Accounting, Finance and Management*, 25(1):18–27, 2018.

[KLS+17] Ágnes Kiss, Jian Liu, Thomas Schneider, N. Asokan, and Benny Pinkas. Private Set Intersection for Unequal Set Sizes with Mobile Applications. *Proceedings on Privacy Enhancing Technologies Symposium (PETS '17)*, 2017(4):177–197, 2017.

[KNR+22] Hakan Kayan, Matthew Nunes, Omer Rana, Pete Burnap, and Charith Perera. Cybersecurity of Industrial Cyber-Physical Systems: A Review. *ACM Computing Surveys*, 54(11s), 2022.

[KNT+21] Ike Kunze, Philipp Niemietz, Liam Tirpitz, René Glebke, Daniel Trauth, Thomas Bergs, and Klaus Wehrle. Detecting Out-Of-Control Sensor Signals in Sheet Metal Forming using In-Network Computing. In *Proceedings of the 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE '21)*. IEEE, 2021.

[Kor20] István Koren. *DevOpsUse: community-driven continuous innovation of web information infrastructures*. PhD thesis, RWTH Aachen University, 2020.

[KOWFC10] Florian Kerschbaum, Nina Oertel, and Leonardo Weiss Ferreira Chaves. Privacy-Preserving Computation of Benchmarks on Item-Level Data Using RFID. In *Proceedings of the 3rd ACM Conference on Wireless Network Security (WiSec '10)*, pages 105–110. ACM, 2010.

[Koz04] Metin Kozak. *Destination Benchmarking: Concepts, Practices and Operations*. CABI, 2004.

[KPW21] Michael Kretschmer, Jan Pennekamp, and Klaus Wehrle. Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. *ACM Transactions on the Web*, 15(4), 2021.

[KS19] Jan Martin Keil and Sirko Schindler. Comparison and evaluation of ontologies for units of measurement. *Semantic Web*, 10(1):33–51, 2019.

[KSSV14] Patrick Koeberl, Steffen Schulz, Ahmad-Reza Sadeghi, and Vijay Varadharajan. TrustLite: A Security Architecture for Tiny Embedded Devices. In *Proceedings of the 9th European Conference on Computer Systems (EuroSys '14)*. ACM, 2014.

[KT19] Florian Kerschbaum and Anselme Tueno. An Efficiently Searchable Encrypted Data Structure for Range Queries. In *Proceedings of the 24th European Symposium on Research in Computer Security (ESORICS '19)*, volume 11736, pages 344–364. Springer, 2019.

[Kus17] Andrew Kusiak. Smart manufacturing must embrace big data. *Nature*, 544(7648):23–25, 2017.

[KVS07] Anne-Marie Kermarrec and Maarten Van Steen. Gossiping in Distributed Systems. *ACM SIGOPS Operating Systems Review*, 41(5):2–7, 2007.

[KWBK+23] Aline Kluge-Wilkes, Ralph Baier, Ike Kunze, Aleksandra Müller, Amir Shahidi, Dominik Wolfschläger, Christian Brecher, Burkhard Corves, Mathias Hüsing, Verena Nitsch, Robert H. Schmitt, and Klaus Wehrle. Modular Control and Services to Operate Lineless Mobile Assembly Systems. In *Internet of Production: Fundamentals, Applications and Proceedings*, pages 303–328. Springer, 2023.

[KWP+22a] Dominik Kus, Eric Wagner, Jan Pennekamp, Konrad Wolsing, Ina Berenice Fink, Markus Dahlmanns, Klaus Wehrle, and Martin Henze. A False Sense of Security? Revisiting the State of Machine Learning-Based Industrial Intrusion Detection. In *Proceedings of the 8th ACM Cyber-Physical System Security Workshop (CPSS '22)*, pages 73–84. ACM, 2022.

[KWP+22b] Dominik Kus, Konrad Wolsing, Jan Pennekamp, Eric Wagner, Martin Henze, and Klaus Wehrle. Poster: Ensemble Learning for Industrial Intrusion Detection. Technical Report RWTH-2022-10809, RWTH Aachen University, 2022. 38th Annual Computer Security Applications Conference (ACSAC '22).

[KZ17] Lynda Kacha and Abdelhafid Zitouni. An Overview on Data Security in Cloud Computing. *Proceedings of the Computational Methods in Systems and Software (CoMeSySo '17)*, 661:250–261, 2017.

[Lai23] Kim Laine. Bootstrapping module in Microsoft Homomorphic Encryption Library SEAL. `https://stackoverflow.com/questions/54920783/bootstrapping-module-in-microsoft-homomorphic-encryption-library-seal`, 2019 (accessed April 4, 2023).

[LAS13] LASER Workshop. The LASER Workshop. `https://laser-workshop.org/`, 2013.

[LBM10] Tom Longstaff, David Balenson, and Mark Matties. Barriers to science in security. In *Proceedings of the 26th Annual Computer Security Applications Conference (ACSAC '10)*, pages 127–129. ACM, 2010.

[LEG99] Douglas M. Lambert, Margaret A. Emmelhainz, and John T. Gardner. Building successful logistics partnerships. *Journal of Business Logistics*, 20(1):165–181, 1999.

[Len22] Stefan Lenz. Classifying methods used in verifiable privacy-preserving computation. Advisor: Jan Pennekamp. Examiner: Klaus Wehrle. Seminar Paper, RWTH Aachen University, 2022.

[Les23] Jacques Leslie. How Climate Change Is Disrupting the Global Supply Chain. `https://e360.yale.edu/features/how-climate-change-is-disrupting-the-global-supply-chain`, 2022 (accessed April 4, 2023).

[LGS17] Prasanth Lade, Rumi Ghosh, and Soundar Srinivasan. Manufacturing Analytics and Industrial Internet of Things. *IEEE Intelligent Systems*, 32(3):74–79, 2017.

[LH21] Yannik Lockner and Christian Hopmann. Induced network-based transfer learning in injection molding for process modelling and optimization with artificial neural networks. *The International Journal of Advanced Manufacturing Technology*, 112(11):3501–3513, 2021.

[LHZ22] Yannik Lockner, Christian Hopmann, and Weibo Zhao. Transfer learning with artificial neural networks between injection molding processes and different polymer materials. *Journal of Manufacturing Processes*, 73:395–408, 2022.

[Lin05] Yehida Lindell. Secure Multiparty Computation for Privacy-Preserving Data Mining. In *Encyclopedia of Data Warehousing and Mining*, chapter 189, pages 1005–1009. IGI Global, 2005.

[LJ20] Martin Liebenberg and Matthias Jarke. Information Systems Engineering with DigitalShadows: Concept and Case Studies. In *Proceedings of the 32nd International Conference on Advanced Information Systems Engineering (CAiSE '20)*, volume 12127, pages 70–84. Springer, 2020.

[LJ23] Martin Liebenberg and Matthias Jarke. Information systems engineering with Digital Shadows: Concept and use cases in the Internet of Production. *Information Systems*, 114, 2023.

[LK23] Matt Leonard and Shefali Kapadia. Timeline: How the Suez Canal blockage unfolded across supply chains. `https://www.supplychaindive.com/news/timeline-ever-given-evergreen-blocked-suez-canal-supply-chain/597660/`, 2021 (accessed April 4, 2023).

[LKG+18] Elena Simona Lohan, Mike Koivisto, Olga Galinina, Sergey Andreev, Antti Tolli, Giuseppe Destino, Mario Costa, Kari Leppanen, Yevgeni Koucheryavy, and Mikko Valkama. Benefits of Positioning-Aided Communication Technology in High-Frequency Industrial IoT. *IEEE Communications Magazine*, 56(12):142–148, 2018.

[LL23] Anja Leckel and Maria Linnartz. Towards The Internet Of Production–How To Increase Data Sharing For Successful Supply Chain Collaboration. *Journal of Production Systems and Logistics*, 3, 2023.

[LM12] Jon Loeliger and Matthew McCullough. *Version Control with Git: Powerful tools and techniques for collaborative software development.* O'Reilly Media, 2012.

[LMS+21] Maria Linnartz, Ursula Motz, Tobias Schröer, Volker Stich, Kai Müller, and Christoph Greb. Increasing Resilience in Procurement in the Context of the Internet of Production. *Journal of Production Systems and Logistics*, 1(2021), 2021.

[LPMW22] Johannes Lohmöller, Jan Pennekamp, Roman Matzutt, and Klaus Wehrle. On the Need for Strong Sovereignty in Data Ecosystems. In *Proceedings of the 1st International Workshop on Data Ecosystems (DEco '22)*, volume 3306. CEUR Workshop Proceedings, 2022.

[LRWW20] Seppo Leminen, Mervi Rajahonka, Robert Wendelin, and Mika Westerlund. Industrial internet of things business models in the machine-to-machine context. *Industrial Marketing Management*, 84:298–311, 2020.

[LSLN+16] Tuan Le, Gabriel Salles-Loustau, Laleh Najafizadeh, Mehdi Javanmard, and Saman Zonouz. Secure Point-of-Care Medical Diagnostics via Trusted Sensing and Cyto-Coded Passwords. In *Proceedings of the 2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN '16)*, pages 583–594. IEEE, 2016.

[LSTS20] Tian Li, Anit Kumar Sahu, Ameet Talwalkar, and Virginia Smith. Federated Learning: Challenges, Methods, and Future Directions. *IEEE Signal Processing Magazine*, 37(3):50–60, 2020.

[LW11] Allison Lewko and Brent Waters. Decentralizing Attribute-Based Encryption. In *Proceedings of the 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT '11)*, volume 6632, pages 568–588. Springer, 2011.

[LWW+23] Olav Lamberts, Konrad Wolsing, Eric Wagner, Jan Pennekamp, Jan Bauer, Klaus Wehrle, and Martin Henze. SoK: Evaluations in Industrial Intrusion Detection Research. *Journal of Systems Research*, 3(1), 2023.

[LXC12] Jing Liu, Yang Xiao, and C. L. Philip Chen. Authentication and Access Control in the Internet of Things. In *Proceeedings of the 2012 32nd International Conference on Distributed Computing Systems Workshops (ICDCSW '12)*, pages 588–592. IEEE, 2012.

[LYZL18] Joseph K. Liu, Tsz Hon Yuen, Peng Zhang, and Kaitai Liang. Time-Based Direct Revocable Ciphertext-Policy Attribute-Based Encryption with Short Revocation List. In *Proceedings of the 16th International Conference on Applied Cryptography and Network Security (ACNS '18)*, volume 10892, pages 516–534. Springer, 2018.

[MAB+13] Frank McKeen, Ilya Alexandrovich, Alex Berenzon, Carlos V. Rozas, Hisham Shafi, Vedvyas Shanbhogue, and Uday R. Savagaonkar. Innovative instructions and software model for isolated execution. In *Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy (HASP '13)*, 2013.

[MAL23] Daniel Morales, Isaac Agudo, and Javier Lopez. Private set intersection: A systematic literature review. *Computer Science Review*, 49, 2023.

[MBLT13] Douglas Maughan, David Balenson, Ulf Lindqvist, and Zachary Tudor. Crossing the "Valley of Death": Transitioning Cybersecurity Research into Practice. *IEEE Security & Privacy*, 11(2):14–23, 2013.

[MBP23] Lukas Moschko, Vera Blazevic, and Frank T. Piller. Paradoxes of implementing digital manufacturing systems: A longitudinal study of digital innovation projects for disruptive change. *Journal of Product Innovation Management*, 40(4):506–529, 2023.

[McC02] Michael McClellan. *Collaborative Manufacturing: Using Real-Time Information to Support the Supply Chain.* CRC Press, 1st edition, 2002.

[MCdD07] Thierry Moyaux, Brahim Chaib-draa, and Sophie D'Amours. Information Sharing as a Coordination Mechanism for Reducing the Bullwhip Effect in a Supply Chain. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 37(3):396–409, 2007.

[MDKJ19] Sidra Malik, Volkan Dedeoglu, Salil S. Kanhere, and Raja Jurdak. TrustChain: Trust Management in Blockchain and IoT supported Supply Chains. In *Proceedings of the 2019 IEEE International Conference on Blockchain (Blockchain '19)*, pages 184–193. IEEE, 2019.

[MdRC12] José Maria Viedma Marti and Maria do Rosário Cabrita. *Entrepreneurial Excellence in the Knowledge Economy: Intellectual Capital Benchmarking Systems.* Palgrave Macmillan, 2012.

[MEE19] Mourad El Maouchi, Oğuzhan Ersoy, and Zekeriya Erkin. DECOUPLES: A Decentralized, Unlinkable and Privacy-preserving Traceability System for the Supply Chain. In *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (SAC '19)*, pages 364–373. ACM, 2019.

[Mer14] Dirk Merkel. Docker: Lightweight Linux Containers for Consistent Development and Deployment. *Linux Journal*, 2014(239), 2014.

[MGDC+17] Pieter Maene, Johannes Götzfried, Ruan De Clercq, Tilo Müller, Felix Freiling, and Ingrid Verbauwhede. Hardware-Based Trusted Computing Architectures for Isolation and Attestation. *IEEE Transactions on Computers*, 67(3):361–374, 2017.

[MGP+21] Simon Mangel, Lars Gleim, Jan Pennekamp, Klaus Wehrle, and Stefan Decker. Data Reliability and Trustworthiness through Digital Transmission Contracts. In *Proceedings of the 18th Extended Semantic Web Conference (ESWC '21)*, volume 12731, pages 265–283. Springer, 2021.

[MH20] Hamid Mozaffari and Amir Houmansadr. Heterogeneous Private Information Retrieval. In *Proceedings of the 28th Annual Network and Distributed System Security Symposium (NDSS '20)*. Internet Society, 2020.

[MHH+18] Roman Matzutt, Jens Hiller, Martin Henze, Jan Henrik Ziegeldorf, Dirk Müllmann, Oliver Hohlfeld, and Klaus Wehrle. A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin. In *Proceedings of the 22th International Conference on Financial Cryptography and Data Security (FC '18)*, volume 10957, pages 420–438. Springer, 2018.

[MHK+18] Mimi Ma, Debiao He, Neeraj Kumar, Kim-Kwang Raymond Choo, and Jianhua Chen. Certificateless Searchable Public Key Encryption Scheme for Industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, 14(2):759–767, 2018.

[MHMW24] Roman Matzutt, Martin Henze, Dirk Müllmann, and Klaus Wehrle. Illicit Blockchain Content – Its Different Shapes, Consequences, and Remedies. In *Blockchains: A Handbook on Fundamentals, Platforms and Applications*, volume 105 of *Advances in Information Security*, pages 301–336. Springer, 2024.

[MHS17] Christopher Millard, W. Kuan Hon, and Jatinder Singh. Internet of Things Ecosystems: Unpacking Legal Relationships and Liabilities. In *Proceedings of the 2017 IEEE International Conference on Cloud Engineering (IC2E '17)*, pages 286–291. IEEE, 2017.

[Mic18] Microsoft, Inc. Microsoft SEAL. `https://github.com/Microsoft/SEAL`, 2018.

[Mic21] Jan-Gustav Michnia. Improving Privacy-Preserving Company Benchmarking with Modern FHE Schemes. Bachelor's thesis, RWTH Aachen University, 2021.

[MKJ18] Sidra Malik, Salil S. Kanhere, and Raja Jurdak. ProductChain: Scalable Blockchain Framework to Support Provenance in Supply Chains. In *Proceedings of the 2018 IEEE 17th International Symposium on Network Computing and Applications (NCA '18)*. IEEE, 2018.

[MKP+21] Roman Matzutt, Benedikt Kalde, Jan Pennekamp, Arthur Drichel, Martin Henze, and Klaus Wehrle. CoinPrune: Shrinking Bitcoin's Blockchain Retrospectively. *IEEE Transactions on Network and Service Management*, 18(3):3064–3078, 2021.

[MMH+15] Jef Maerien, Sam Michiels, Danny Hughes, Christophe Huygens, and Wouter Joosen. SecLooCI: A comprehensive security middleware architecture for shared wireless sensor networks. *Ad Hoc Networks*, 25(Part A):141–169, 2015.

[MMZ+17] Roman Matzutt, Dirk Müllmann, Eva-Maria Zeissig, Christiane Horst, Kai Kasugai, Sean Lidynia, Simon Wieninger, Jan Henrik Ziegeldorf, Gerhard Gudergan, Indra Spiecker gen. Döhmann, Klaus Wehrle, and Martina Ziefle. myneData: Towards a Trusted and User-controlled Ecosystem for Sharing Personal Data. In *INFORMATIK*, volume 275, pages 1073–1084. Gesellschaft für Informatik, 2017.

[Mob17] MobileCoin. MobileCoin - Safe & easy payments at light-speed. `https://mobilecoin.com/`, 2017.

[Mom02] Jesper Momme. Framework for outsourcing manufacturing: strategic and operational implications. *Computers in Industry*, 49(1):59–75, 2002.

[Mon09] MongoDB Inc. MongoDB. `https://www.mongodb.com`, 2009.

[Mos18] Michele Mosca. Cybersecurity in an Era with Quantum Computers: Will We Be Ready? *IEEE Security & Privacy*, 16(5):38–41, 2018.

[Mou22] Dimitris Mourtzis. *Design and Operation of Production Networks for Mass Personalization in the Era of Cloud Technology*. Elsevier, 1st edition, 2022.

[MPBW20] Roman Matzutt, Jan Pennekamp, Erik Buchholz, and Klaus Wehrle. Utilizing Public Blockchains for the Sybil-Resistant Bootstrapping of Distributed Anonymity Services. In *Proceedings of the 15th ACM ASIA Conference on Computer and Communications Security (ASIACCS '20)*, pages 531–542. ACM, 2020.

[MPFS22] Jämes Ménétrey, Marcelo Pasin, Pascal Felber, and Valerio Schiavoni. WaTZ: A Trusted WebAssembly Runtime Environment with Remote Attestation for TrustZone. In *Proceedings of the 2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS '22)*, pages 1177–1189. IEEE, 2022.

[MPW20] Roman Matzutt, Jan Pennekamp, and Klaus Wehrle. A Secure and Practical Decentralized Ecosystem for Shareable Education Material. In *Proceedings of the 34th International Conference on Information Networking (ICOIN '20)*, pages 529–534. IEEE, 2020.

[MPW23] Roman Matzutt, Jan Pennekamp, and Klaus Wehrle. Poster: Accountable Processing of Reported Street Problems. In *Proceedings of the 30th ACM SIGSAC Conference on Computer and Communications Security (CCS '23)*, pages 3591–3593. ACM, 2023.

[MR91] Silvio Micali and Phillip Rogaway. Secure Computation. In *Proceedings of the 11th Annual International Cryptology Conference (CRPYTO '91)*, volume 576, pages 392–404. Springer, 1991.

[MR04] Jelena Mirkovic and Peter Reiher. A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM SIGCOMM Computer Communication Review*, 34(2):39–53, 2004.

[MSBAU22a] Aintzane Mosteiro-Sanchez, Marc Barcelo, Jasone Astorga, and Aitor Urbieta. End to End Secure Data Exchange in Value Chains with Dynamic Policy Updates. arXiv:2201.06335, 2022.

[MSBAU22b] Aintzane Mosteiro-Sanchez, Marc Barcelo, Jasone Astorga, and Aitor Urbieta. Too Many Options: A Survey of ABE Libraries for Developers. arXiv:2209.12742, 2022.

[MSCD18] Giovanni Miragliotta, Andrea Sianesi, Elisa Convertini, and Rossella Distante. Data driven management in Industry 4.0: a method to measure Data Productivity. *IFAC-PapersOnLine*, 51(11):19–24, 2018. Proceedings of the 16th IFAC Symposium on Information Control Problems in Manufacturing (INCOM '18).

[MSS+21] Praveen Kumar Malik, Rohit Sharma, Rajesh Singh, Anita Gehlot, Suresh Chandra Satapathy, Waleed S. Alnumay, Danilo Pelusi, Uttam Ghosh, and Janmenjoy Nayak. Industrial Internet of Things and its Applications in Industry 4.0: State of The Art. *Computer Communications*, 166:125–139, 2021.

[MST+14] Hrelja Marko, Klancnik Simon, Irgolic Tomaz, Paulic Matej, Balic Joze, and Brezocnik Miran. Turning Parameters Optimization Using Particle Swarm Optimization. *Procedia Engineering*, 69:670–677, 2014.

[MTT+18] Richard Meyes, Hasan Tercan, Thomas Thiele, Alexander Krämer, Julian Heinisch, Martin Liebenberg, Gerhard Hirt, Christian Hopmann, Gerhard Lakemeyer, Tobias Meisen, and Sabina Jeschke. Interdisciplinary Data Driven Production Process Analysis for the Internet of Production. *Procedia Manufacturing*, 26:1065–1076, 2018.

[MVS13] Mohammad Saleh Meiabadi, Abbas Vafaeesefat, and Fatemeh Sharifi. Optimization of plastic injection molding process by combination of artificial neural network and genetic algorithm. *Journal of Optimization in Industrial Engineering*, 6(13):49–54, 2013.

[MWM+16] David C. Mills, Kathy Wang, Brendan Malone, Anjana Ravi, Jeffrey Marquardt, Anton I. Badev, Timothy Brezinski, Linda Fahy, Kimberley Liao, Vanessa Kargenian, Max Ellithorpe, Wendy Ng, and Maria Baird. Distributed Ledger Technology in Payments, Clearing, and Settlement. Technical report, 2016.

[MWS+19] Sinisa Matetic, Karl Wüst, Moritz Schneider, Kari Kostiainen, Ghassan Karame, and Srdjan Capkun. BITE: Bitcoin Lightweight Client Privacy using Trusted Execution. In *Proceedings of the 28th USENIX Security Symposium (SEC '19)*, pages 783–800. USENIX Association, 2019.

[NBF+16] Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, and Steven Goldfeder. *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*. Princeton University Press, 2016.

[NBM+17] Job Noorman, Jo Van Bulck, Jan Tobias Mühlberg, Frank Piessens, Pieter Maene, Bart Preneel, Ingrid Verbauwhede, Johannes Götzfried, Tilo Müller, and Felix Freiling. Sancus 2.0: A Low-Cost Security Architecture for IoT Devices. *ACM Transactions on Privacy and Security*, 20(3), 2017.

[NER+19] Ivan De Oliveira Nunes, Karim Eldefrawy, Norrathep Rattanavipanon, Michael Steiner, and Gene Tsudik. VRASED: A Verified Hardware/Software Co-Design for Remote Attestation. In *Proceedings of the 28th USENIX Security Symposium (SEC '19)*, pages 1429–1446. USENIX Association, 2019.

[NKU+20] Philipp Niemietz, Tobias Kaufmann, Martin Unterberg, Daniel Trauth, and Thomas Bergs. Towards an Adaptive Production Chain for Sustainable Sheet-Metal Blanked Components. In *Proceedings of the 10th Congress of the German Academic Association for Production Technology (WGP '20)*, pages 34–44. Springer, 2020.

[NLO15] Kim Thuat Nguyen, Maryline Laurent, and Nouha Oualha. Survey on secure communication protocols for the Internet of Things. *Ad Hoc Networks*, 32:17–31, 2015.

[NLV11] Michael Naehrig, Kristin Lauter, and Vinod Vaikuntanathan. Can homomorphic encryption be practical? In *Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop (CCSW '11)*, pages 113–124. ACM, 2011.

[NMP17] Job Noorman, Jan Tobias Mühlberg, and Frank Piessens. Authentic Execution of Distributed Event-Driven Applications with a Small TCB. In *Proceedings of the 13th International Workshop on Security and Trust Management (STM '17)*, volume 10547, pages 55–71. Springer, 2017.

[NN17] Rajesh Narang and Tanmay Narang. Preserving Confidentiality and Privacy of Sensitive Data in e-Procurement System. *International Journal of Cyber-Security and Digital Forensics*, 6(4):186–197, 2017.

[NPP01] Moni Naor, Benny Pinkas, and Benny Pinkas. Efficient Oblivious Transfer Protocols. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '01)*, pages 448–457. SIAM, 2001.

[NPS99] Moni Naor, Benny Pinkas, and Reuban Sumner. Privacy Preserving Auctions and Mechanism Design. In *Proceedings of the 1st ACM Conference on Electronic Commerce (EC '99)*, pages 129–139. ACM, 1999.

[NS91] Robert A. Novack and Stephen W. Simco. The Industrial Procurement Process: A Supply Chain Perspective. *Journal of Business Logistics*, 12(1):145–168, 1991.

[NT00] Khanh Quoc Nguyen and Jacques Traoré. An Online Public Auction Protocol Protecting Bidder Privacy. In *Proceedings of the 5th Australasian Conference on Information Security and Privacy (ACISP '00)*, volume 1841, pages 427–442. Springer, 2000.

[NWL10] Gilbert N. Nyaga, Judith M. Whipple, and Daniel F. Lynch. Examining supply chain relationships: Do buyer and supplier perspectives on collaborative relationships differ? *Journal of Operations Management*, 28(2):101–114, 2010.

[OAC+16] Boris Otto, Sören Auer, Jan Cirullies, Jan Jürjens, Nadja Menz, Jochen Schon, and Sven Wenzel. Industrial Data Space: Digital Souvereignity over Data. White paper, Fraunhofer, 2016.

[OBNB23] Lucia Ortjohann, Marco Becker, Philipp Niemietz, and Thomas Bergs. Monitoring of fluctuating material properties for optimizing sheet-metal forming processes: a systematic literature review. In *Proceedings of the 26th International ESAFORM Conference on Material Forming (ESAFORM '23)*, volume 28, pages 2071–2080. Materials Research Forum, 2023.

[ODJ+22] Ilhaam A. Omar, Mazin Debe, Raja Jayaraman, Khaled Salah, Mohammad Omar, and Junaid Arshad. Blockchain-based Supply Chain Traceability for COVID-19 personal protective equipment. *Computers & Industrial Engineering*, 167, 2022.

[OECD13a] Organisation for Economic Co-operation and Development. Fighting bid rigging in public procurement. `https://www.oecd.org/competition/cartels/fightingbidriggininpublicprocurement.htm`, 2013 (accessed April 4, 2023).

[OECD13b] Organisation for Economic Co-operation and Development. Cartels and anti-competitive agreements. `https://www.oecd.org/competition/cartels/`, 2013 (accessed March 6, 2022).

[OJ19] Boris Otto and Matthias Jarke. Designing a multi-sided data platform: findings from the International Data Spaces case. *Electronic Markets*, 29(4):561–580, 2019.

[Ope22] OpenFHE. OpenFHE.org. `https://www.openfhe.org/`, 2022.

[OTG07] Tim A. Osswald, Lih-Sheng Turng, and Paul J. Gramann. *Injection Molding Handbook*. Carl Hanser, 2nd edition, 2007.

[PAAD18] Stephen Pollard, Guy Adams, Faisal Azhar, and Fraser Dickin. Authentication of 3D Printed Parts using 3D Physical Signatures. In *Proceedings of the NIP & Digital Fabrication Conference, Printing for Fabrication 2018*, pages 196–201. Society for Imaging Science and Technology, 2018.

[PAB+24] Jan Pennekamp, Fritz Alder, Lennart Bader, Gianluca Scopelliti, Klaus Wehrle, and Jan Tobias Mühlberg. Securing Sensing in Supply Chains: Opportunities, Building Blocks, and Designs. *IEEE Access*, 12:9350–9368, 2024.

[Pai99] Pascal Paillier. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT '99)*, volume 1592, pages 223–238. Springer, 1999.

[PAM+20] Jan Pennekamp, Fritz Alder, Roman Matzutt, Jan Tobias Mühlberg, Frank Piessens, and Klaus Wehrle. Secure End-to-End Sensing in Supply Chains. In *Proceedings of the 2020 IEEE Conference on Communications and Network Security (CNS '20)*. IEEE, 2020. Proceedings of the 5th International Workshop on Cyber-Physical Systems Security (CPS-Sec '20).

[Par15] David Parmenter. *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs*. Wiley, 3rd edition, 2015.

[Pat17] Constantinos Patsakis. OrderRevealingEncryption. `https://github.com/kpatsakis/OrderRevealingEncryption`, 2017.

[PBB+23] Jan Pennekamp, Anastasiia Belova, Thomas Bergs, Matthias Bodenbenner, Andreas Bührig-Polaczek, Markus Dahlmanns, Ike Kunze, Moritz Kröger, Sandra Geisler, Martin Henze, Daniel Lütticke, Benjamin Montavon, Philipp Niemietz, Lucia Ortjohann, Maximilian Rudack, Robert H. Schmitt, Uwe Vroomen, Klaus Wehrle, and Michael Zeng. Evolving the Digital Industrial Infrastructure for Production: Steps Taken and the Road Ahead. In *Internet of Production: Fundamentals, Applications and Proceedings*, Interdisciplinary Excellence Accelerator Series, pages 35–60. Springer, 2023.

[PBD+21] Jan Pennekamp, Erik Buchholz, Markus Dahlmanns, Ike Kunze, Stefan Braun, Eric Wagner, Matthias Brockmann, Klaus Wehrle, and Martin Henze. Collaboration is not Evil: A Systematic Look at Security Research for Industrial Use. In *Proceedings of the Workshop on Learning from Authoritative Security Experiment Results (LASER '20)*. ACSAC, 2021.

[PBL+20] Jan Pennekamp, Erik Buchholz, Yannik Lockner, Markus Dahlmanns, Tiandong Xi, Marcel Fey, Christian Brecher, Christian Hopmann, and Klaus Wehrle. Privacy-Preserving Production Process Parameter Exchange. In *Proceedings of the 36th Annual Computer Security Applications Conference (ACSAC '20)*, pages 510–525. ACM, 2020.

[PBM+20] Jan Pennekamp, Lennart Bader, Roman Matzutt, Philipp Niemietz, Daniel Trauth, Martin Henze, Thomas Bergs, and Klaus Wehrle. Private Multi-Hop Accountability for Supply Chains. In *Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops '20)*. IEEE, 2020. Proceedings of the 1st Workshop on Blockchain for IoT and Cyber-Physical Systems (BIoTCPS '20).

[PBW+24] Jan Pennekamp, Lennart Bader, Eric Wagner, Jens Hiller, Roman Matzutt, and Klaus Wehrle. Blockchain Technology Accelerating Industry 4.0. In *Blockchains: A Handbook on Fundamentals, Platforms and Applications*, volume 105 of *Advances in Information Security*, pages 531–564. Springer, 2024.

[PDF+23] Jan Pennekamp, Markus Dahlmanns, Frederik Fuhrmann, Timo Heutmann, Alexander Kreppein, Dennis Grunert, Christoph Lange, Robert H. Schmitt, and Klaus Wehrle. Offering Two-Way Privacy for Evolved Purchase Inquiries. *ACM Transactions on Internet Technology*, 23(4), 2023.

[PDG+19] Jan Pennekamp, Markus Dahlmanns, Lars Gleim, Stefan Decker, and Klaus Wehrle. Security Considerations for Collaborations in an Industrial IoT-based Lab of Labs. In *Proceedings of the 3rd IEEE Global Conference on Internet of Things (GCIoT '19)*. IEEE, 2019.

[Ped92] Torben Pryds Pedersen. Non-Interactive and Information-Theoretic Secure Verifiable Secret Sharing. In *Proceedings of the 11th Annual International Cryptology Conference (CRPYTO '91)*, volume 576, pages 129–140. Springer, 1992.

[Pen24] Jan Pennekamp. Evolving the Industrial Internet of Things: The Advent of Secure Collaborations. In *Proceedings of the 2024 IEEE/IFIP Network Operations and Management Symposium (NOMS '24)*. IEEE, 2024. In Press.

[Per22] PerFail 2022. PerFail 2022 – First International Workshop on Negative Results in Pervasive Computing. `https://perfail-workshop.github.io/2022/`, 2022.

[PFCN22] Warwick Powell, Marcus Foth, Shoufeng Cao, and Valéri Natanelov. Garbage in garbage out: The precarious link between IoT and blockchain in food supply chains. *Journal of Industrial Information Integration*, 25, 2022.

[PGH+19] Jan Pennekamp, René Glebke, Martin Henze, Tobias Meisen, Christoph Quix, Rihan Hai, Lars Gleim, Philipp Niemietz, Maximilian Rudack, Simon Knape, Alexander Epple, Daniel Trauth, Uwe Vroomen, Thomas Bergs, Christian Brecher, Andreas Bührig-Polaczek, Matthias Jarke, and Klaus Wehrle. Towards an Infrastructure Enabling the Internet of Production. In *Proceedings of the 2nd IEEE International Conference on Industrial Cyber Physical Systems (ICPS '19)*, pages 31–37. IEEE, 2019.

[PGP+17] Sandro Pinto, Tiago Gomes, Jorge Pereira, Jorge Cabral, and Adriano Tavares. IIoTEED: An Enhanced, Trusted Execution Environment for Industrial IoT Edge Devices. *IEEE Internet Computing*, 21(1):40–47, 2017.

[PH10] Andreas Pfitzmann and Marit Hansen. A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management. Technical report, TU Dresden, 2010.

[PHB06] Alexander Pretschner, Manuel Hilty, and David Basin. Distributed usage control. In *Communications of the ACM*, volume 49, pages 39–44. ACM, 2006.

[PHS+19] Jan Pennekamp, Martin Henze, Simo Schmidt, Philipp Niemietz, Marcel Fey, Daniel Trauth, Thomas Bergs, Christian Brecher, and Klaus Wehrle. Dataflow Challenges in an *Internet* of Production: A Security & Privacy Perspective. In *Proceedings of the ACM Workshop on Cyber-Physical Systems Security & Privacy (CPS-SPC '19)*, pages 27–38. ACM, 2019.

[PHW21] Jan Pennekamp, Martin Henze, and Klaus Wehrle. Unlocking Secure Industrial Collaborations through Privacy-Preserving Computation. *ERCIM News*, 126:24–25, 2021.

[Pil16] Marc Pilkington. Blockchain technology: principles and applications. In *Research Handbook on Digital Transformations*, chapter 11, pages 225–253. Edward Elgar Publishing, 1st edition, 2016.

[PIZ+20] Michael Paik, Jerónimo Irazábal, Dennis Zimmer, Michele Meloni, and Valentin Padurean. immudb: A Lightweight, Performant Immutable Database. Technical report, CodeNotary, 2020.

[PKM18] Abhijeet C Panchal, Vijay M. Khadse, and Parikshit N. Mahalle. Security Issues in IIoT: A Comprehensive Survey of Attacks on IIoT and Its Countermeasures. In *Proceedings of the 2018 IEEE Global Conference on Wireless Computing and Networking (GCWCN '18)*, pages 124–130. IEEE, 2018.

[Pla19] PlasticsEurope. Geschäftsbericht 2018 [annual report 2018]. Technical report, PlasticsEurope Deutschland e.V., 2019.

[PLV+23] Jan Pennekamp, Johannes Lohmöller, Eduard Vlad, Joscha Loos, Niklas Rodemann, Patrick Sapel, Ina Berenice Fink, Setz Schmitz, Christian Hopmann, Matthias Jarke, Günther Schuh, Klaus Wehrle, and Martin Henze. Designing Secure and Privacy-Preserving Information Systems for Industry Benchmarking. In *Proceedings of the 35th International Conference on Advanced Information Systems Engineering (CAiSE '23)*, pages 489–505. Springer, 2023.

[PMK+21] Jan Pennekamp, Roman Matzutt, Salil S. Kanhere, Jens Hiller, and Klaus Wehrle. The Road to Accountable and Dependable Manufacturing. *Automation*, 2(3):202–219, 2021.

[PMK+24] Jan Pennekamp, Roman Matzutt, Christopher Klinkmüller, Lennart Bader, Martin Serror, Eric Wagner, Sidra Malik, Maria Spiß, Jessica Rahn, Tan Gürpinar, Eduard Vlad, Sander J. J. Leemans, Salil S. Kanhere, Volker Stich, and Klaus Wehrle. An Interdisciplinary Survey on Information Flows in Supply Chains. *ACM Computing Surveys*, 56(2), 2024.

[PNT99] H. S. C. Perera, Nagen Nagarur, and Mario T. Tabucanon. Component part standardization: A way to reduce the life-cycle costs of products. *International Journal of Production Economics*, 60–61:109–116, 1999.

[PPW08] Vijayakrishnan Pasupathinathan, Josef Pieprzyk, and Huaxiong Wang. A Fair E-Tendering Protocol. In *Proceedings of the 5th International Conference on Security and Cryptography (SECRYPT '08)*, pages 294–299. SCITEPRESS, 2008.

[Pra16] Sinha Prashant. pybloomfiltermmap3. `https://github.com/prashnts/pybloomfiltermmap3`, 2016.

[PS17] Radim Polčák and Dan Jerker B. Svantesson. *Information Sovereignty: Data Privacy, Sovereign Powers and the Rule of Law*. Edward Elgar Publishing, 1st edition, 2017.

[PS19] Sandro Pinto and Nuno Santos. Demystifying Arm TrustZone: A Comprehensive Survey. *ACM Computing Surveys*, 51(6), 2019.

[PSF+20] Jan Pennekamp, Patrick Sapel, Ina Berenice Fink, Simon Wagner, Sebastian Reuter, Christian Hopmann, Klaus Wehrle, and Martin Henze. Revisiting the Privacy Needs of Real-World Applicable Company Benchmarking. In *Proceedings of the 8th Workshop on Encrypted Computing & Applied Homomorphic Cryptography (WAHC '20)*, pages 31–44. HomomorphicEncryption.org, 2020.

[PSF+23] Sebastian Pütz, Alexander Schollemann, Annika Laura Felter, Mirlinda Hajdari, Alexander Mertens, Seth Schmitz, Günther Schuh, and Verena Nitsch. Towards Human-Centered Best Practice Sharing in Global Production Networks for More Socially Sustainable Production Processes. In *Proceedings of the 2023 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC '23)*. IEEE, 2023.

[PSZ14] Benny Pinkas, Thomas Schneider, and Michael Zohner. Faster Private Set Intersection Based on OT Extension. In *Proceedings of the 23rd USENIX Conference on Security Symposium (SEC '14)*, pages 797–812. USENIX Association, 2014.

[PTM+21] Shenle Pan, Damien Trentesaux, Duncan McFarlane, Benoit Montreuil, Eric Ballot, and George Q. Huang. Digital interoperability in logistics and supply chain management: state-of-the-art and research avenues towards Physical Internet. *Computers in Industry*, 128, 2021.

[PV20] Miguel Pincheira and Massimo Vecchio. Towards Trusted Data on Decentralized IoT Applications: Integrating Blockchain in Constrained Devices. In *Proceedings of the 2020 IEEE International Conference on Communications Workshops (ICC Workshops '20)*. IEEE, 2020. Proceedings of the 1st Workshop on Blockchain for IoT and Cyber-Physical Systems (BIoTCPS '20).

[PWB+19] Peter Poschmann, Manuel Weinke, Andreas Balster, Frank Straube, Hanno Friedrich, and André Ludwig. Realization of ETA Predictions for Intermodal Logistics Networks Using Artificial Intelligence. In *Proceedings of the 4th Interdisciplinary Conference on Production, Logistics and Traffic (ICPLT '19)*, pages 155–176. Springer, 2019.

[QCZ+20] Tie Qiu, Jiancheng Chi, Xiaobo Zhou, Zhaolong Ning, Mohammed Atiquzzaman, and Dapeng Oliver Wu. Edge Computing in Industrial Internet of Things: Architecture, Advances and Challenges. *IEEE Communications Surveys & Tutorials*, 22(4):2462–2488, 2020.

[RAA+19] Mark Russinovich, Edward Ashton, Christine Avanessians, Miguel Castro, Amaury Chamayou, Sylvan Clebsch, Manuel Costa, Cédric Fournet, Matthew Kerner, Sid Krishna, Julien Maffre, Thomas Moscibroda, Kartik Nayak, Olga Ohrimenko, Felix Schuster, Roy Schuster, Alex Shamis, Olga Vrousgou, and Christoph M. Wintersteiger. CCF: A Framework for Building Confidential Verifiable Replicated Services. Technical report, Microsoft, 2019.

[Rab05] Michael O. Rabin. How To Exchange Secrets with Oblivious Transfer. Cryptology ePrint Archive 2005/187, 2005.

[RAB+23] Adrian Karl Rüppel, Muzaffer Ay, Benedikt Biernat, Ike Kunze, Markus Landwehr, Samuel Mann, Jan Pennekamp, Pascal Rabe, Mark P. Sanders, Dominik Scheurenberg, Sven Schiller, Tiandong Xi, Dirk Abel, Thomas Bergs, Christian Brecher, Uwe Reisgen, Robert H. Schmitt, and Klaus Wehrle. Model-Based Controlling Approaches for Manufacturing Processes. In *Internet of Production: Fundamentals, Applications and Proceedings*, Interdisciplinary Excellence Accelerator Series, pages 221–246. Springer, 2023.

[RB16] Sharvari Rautmare and Deepashree M. Bhalerao. MySQL and NoSQL database comparison for IoT application. In *Proceedings of the 2016 IEEE International Conference on Advances in Computer Applications (ICACA '16)*, pages 235–238. IEEE, 2016.

[RDBI20] Víctor Julio Ramírez-Durán, Idoia Berges, and Arantza Illarramendi. ExtruOnt: An ontology for describing a type of manufacturing machine for Industry 4.0 systems. *Semantic Web*, 11(6):887–909, 2020.

[RDF+20] Linus Roepert, Markus Dahlmanns, Ina Berenice Fink, Jan Pennekamp, and Martin Henze. Assessing the Security of OPC UA Deployments. In *Proceedings of the 1st ITG Workshop on IT Security (ITSec '20)*. University of Tübingen, 2020.

[ReZIB21] Wajiha Rehman, Hijab e Zainab, Jaweria Imran, and Narmeen Zakaria Bawany. NFTs: Applications and Challenges. In *Proceedings of the 2021 22nd International Arab Conference on Information Technology (ACIT '21)*. IEEE, 2021.

[RFJ18] Morrakot Raweewan and William G. Ferrell Jr. Information sharing in supply chain collaboration. *Computers & Industrial Engineering*, 126:269–281, 2018.

[RHS23] Ghadafi M. Razak, Linda C. Hendry, and Mark Stevenson. Supply chain traceability: A review of the benefits and its relationship with supply chain resilience. *Production Planning & Control*, 34(11):1114–1134, 2023.

[Rin16a] Peter Rindal. libOTe: an efficient, portable, and easy to use Oblivious Transfer Library. `https://github.com/osu-crypto/libOTe`, 2016.

[Rin16b] Peter Rindal. libPSI: A Private Set Intersection Library. `https://github.com/osu-crypto/libPSI`, 2016.

[Rit18] Margarete Rittstieg. *Einflussfaktoren der Leistungsfähigkeit von Produktionsstandorten in globalen Produktionsnetzwerken [Factors influencing the performance of production sites in global production networks]*. PhD thesis, 2018.

[RKC24] Sushmita Ruj, Salil Kanhere, and Mauro Conti. *Blockchains: A Handbook on Fundamentals, Platforms and Applications*, volume 105 of *Advances in Information Security*. Springer, 1st edition, 2024.

[RKKV17] Jothi Rangasamy, Lakshmi Kuppusamy, Gopi Krishnan, and Velmurugan. Evaluation of puzzle-enabled proxy-assisted denial-of-service protection for web services. *International Journal of Information and Computer Security*, 9(1–2):114–129, 2017.

[Ron10] Armin Ronacher. Flask. `https://palletsprojects.com/p/flask/`, 2010.

[RR17] Peter Rindal and Mike Rosulek. Malicious-Secure Private Set Intersection via Dual Execution. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security (CCS '17)*, pages 1229–1242. ACM, 2017.

[RSA78] Ronald Rivest, Adi Shamir, and Leonard Adleman. A Method for Obtaining Digital Signatures and Public-key Cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[RSS+16] Jan Rüth, Florian Schmidt, Martin Serror, Klaus Wehrle, and Torsten Zimmermann. Communication and Networking for the Industrial Internet of Things. In *Industrial Internet of Things: Cybermanufacturing Systems*, pages 317–346. Springer, 2016.

[Rus23] Mark Russinovich. Azure confidential computing. `https://azure.microsoft.com/en-us/blog/azure-confidential-computing/`, 2018 (accessed April 4, 2023).

[Rya14] Mark D. Ryan. Enhanced Certificate Transparency and End-to-end Encrypted Mail. In *Proceedings of the 21st Annual Network and Distributed System Security Symposium (NDSS '14)*. Internet Society, 2014.

[SAC19] Christophe Schinckus, Mohammadreza Akbari, and Steve Clarke. Corporate Social Responsibility in Sustainable Supply Chain Management: An Econo-Bibliometric Perspective. *Theoretical Economics Letters*, 9(1):247–270, 2019.

[SAE23] Tarek Sultan Al Essa. 5 ways the COVID-19 pandemic has changed the supply chain. `https://www.weforum.org/agenda/2022/01/5-ways-the-covid-19-pandemic-has-changed-the-supply-chain/`, 2022 (accessed April 4, 2023).

[Sah03] B. S. Sahay. Supply chain collaboration: the key to value creation. *Work Study*, 52(2):76–83, 2003.

[San09] Salvatore Sanfilippo. Redis. `https://redis.io/`, 2009.

[SC08] Leo Sauermann and Richard Cyganiak. Cool URIs for the Semantic Web. W3C Interest Group Note, 2008.

[SCC+17] Anatoly P. Sobolev, Simone Circi, Donatella Capitani, Cinzia Ingallina, and Luisa Mannina. Molecular fingerprinting of food authenticity. *Current Opinion in Food Science*, 16:59–66, 2017.

[SCDF16] Jordi Soria-Comas and Josep Domingo-Ferrer. Big Data Privacy: Challenges to Privacy Principles and Models. *Data Science and Engineering*, 1(1):21–28, 2016.

[Sch16] Christian Schröder. The Challenges of Industry 4.0 for Small and Medium-sized Enterprises. Technical report, Friedrich-Ebert-Stiftung, 2016.

[SCL+18] Wencheng Sun, Zhiping Cai, Yangyang Li, Fang Liu, Shengqun Fang, and Guoyan Wang. Security and Privacy in the Medical Internet of Things: A Review. *Security and Communication Networks*, 2018, 2018.

[SCR+11] Elaine Shi, T-H. Hubert Chan, Eleanor Rieffel, Richard Chow, and Dawn Song. Privacy-Preserving Aggregation of Time-Series Data. In *Proceedings of the 18th Network and Distributed System Security Symposium (NDSS '11)*. Internet Society, 2011.

[SCS+21] João B. F. Sequeiros, Francisco T. Chimuco, Musa G. Samaila, Mário M. Freire, and Pedro R. M. Inácio. Attack and System Modeling Applied to IoT, Cloud, and Mobile Ecosystems: Embedding Security by Design. *ACM Computing Surveys*, 53(2), 2021.

[SGM06] Patricia M. Swafford, Soumen Ghosh, and Nagesh Murthy. The antecedents of supply chain agility of a firm: Scale development and model testing. *Journal of Operations Management*, 24(2):170–188, 2006.

[SH13] Ingo Schmidt and Justus Haucap. *Wettbewerbspolitik und Kartellrecht: Eine interdisziplinäre Einführung [Competition Policy and Cartel Law: An Interdisciplinary Introduction]*. De Gruyter, 10th edition, 2013.

[Sha79] Adi Shamir. How to Share a Secret. *Communications of the ACM*, 22(11):612–613, 1979.

[SHH+21] Martin Serror, Sacha Hack, Martin Henze, Marko Schuba, and Klaus Wehrle. Challenges and Opportunities in Securing the Industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, 17(5):2985–2996, 2021.

[Siu20] Alexander Stanislaw David Siuda. Web-Based Privacy-Preserving Comparison of KPIs. Bachelor's thesis, RWTH Aachen University, 2020.

[SJK+19] Le Hoang Son, Sudan Jha, Raghvendra Kumar, Jyotir Moy Chatterjee, and Manju Khari. Collaborative handshaking approaches between internet of computing and internet of things towards a smart world: a review from 2009–2017. *Telecommunication Systems*, 70(4):617–634, 2019.

[SKEEA18] Cetin Sahin, Brandon Kuczenski, Omer Egecioglu, and Amr El Abbadi. Privacy-Preserving Certification of Sustainability Metrics. In *Proceedings of the 8th ACM Conference on Data and Application Security and Privacy (CODASPY '18)*, pages 53–63. ACM, 2018.

[SKPP23] Colin Schulz, Sebastian Kortmann, Frank T. Piller, and Patrick Pollok. Growing with smart products: Why customization capabilities matter for manufacturing firms. *Journal of Product Innovation Management*, 40(6):794–816, 2023.

[SKTG23] Tarek Stolz, István Koren, Liam Tirpitz, and Sandra Geisler. GALOIS: A Hybrid and Platform-Agnostic Stream Processing Architecture. In *Proceedings of the International Workshop on Big Data in Emergent Distributed Environments (BiDEDE '23)*. ACM, 2023.

[SLH+17] João Sá Sousa, Cédric Lefebvre, Zhicong Huang, Jean Louis Raisaro, Carlos Aguilar-Melchor, Marc-Olivier Killijian, and Jean-Pierre Hubaux. Efficient and secure outsourcing of genomic data storage. *BMC Medical Genomics*, 10 (Suppl 2), 2017.

[SMF18] Somayeh Sobati-Moghadam and Amjad Fayoumi. Private Collaborative Business Benchmarking in the Cloud. In *Proceedings of the Science and Information Conference (SAI '18)*, volume 857, pages 1359–1365. Springer, 2018.

[SMGMM22] Kaja Schmidt, Gonzalo Munilla Garrido, Alexander Mühle, and Christoph Meinel. Mitigating Sovereign Data Exchange Challenges: A Mapping to Apply Privacy- and Authenticity-Enhancing Technologies. In *Proceedings of the 19th International Conference on Trust and Privacy in Digital Business (TrustBus '22)*, volume 13582, pages 50–65. Springer, 2022.

[SMS+22] Moritz Schneider, Ramya Jayaram Masti, Shweta Shinde, Srdjan Capkun, and Ronald Perez. SoK: Hardware-supported Trusted Execution Environments. arXiv:2205.12742, 2022.

[SN19] Ashutosh Sheel and Vishnu Nath. Effect of blockchain technology adoption on supply chain adaptability, agility, alignment and performance. *Management Research Review*, 42(12):1353–1374, 2019.

[Sol09] Ask Solem. Celery: Distributed Task Queue. `http://www.celeryproject.org/`, 2009.

[SOM14] Fadi Shrouf, Joaquin Ordieres, and Giovanni Miragliotta. Smart factories in Industry 4.0: A review of the concept and of energy management approached in production based on the Internet of Things paradigm. In *Proceedings of the 2014 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM '14)*, pages 697–701. IEEE, 2014.

[Spi12] Sarah Spiekermann. The Challenges of Privacy by Design. *Communications of the ACM*, 55(7):38–40, 2012.

[SPL+15] Keith Stouffer, Victoria Pillitteri, Suzanne Lightman, Marshall Abrams, and Adam Hahn. Guide to Industrial Control Systems (ICS) Security. NIST SP 800-82 Rev. 2, 2015.

[SPN+23] Gianluca Scopelliti, Sepideh Pouyanrad, Job Noorman, Fritz Alder, Christoph Baumann, Frank Piessens, and Jan Tobias Mühlberg. End-to-End Security for Distributed Event-Driven Enclave Applications on Heterogeneous TEEs. *ACM Transactions on Privacy and Security*, 26(3), 2023.

[SQL00] SQLite. SQLite. `https://www.sqlite.org/`, 2000.

[SrcC20] Jan Pennekamp, Erik Buchholz, Yannik Lockner, Markus Dahlmanns, Tiandong Xi, Marcel Fey, Christian Brecher, Christian Hopmann, and Klaus Wehrle. Privacy-Preserving Production Process Parameter Exchange. `https://github.com/COMSYS/parameter-exchange`, 2020.

[SrcC21] Gianluca Scopelliti, Sepideh Pouyanrad, Job Noorman, Fritz Alder, Christoph Baumann, Frank Piessens, and Jan Tobias Mühlberg. Authentic Execution Framework. `https://github.com/AuthenticExecution/.github`, 2021.

[SrcC23a] Jan Pennekamp, Markus Dahlmanns, Frederik Fuhrmann, Timo Heutmann, Alexander Kreppein, Dennis Grunert, Christoph Lange, Robert H. Schmitt, and Klaus Wehrle. Offering Two-Way Privacy for Evolved Purchase Inquiries. `https://github.com/COMSYS/purchase-inquiries`, 2023.

[SrcC23b] Jan Pennekamp, Johannes Lohmöller, Eduard Vlad, Joscha Loos, Niklas Rodemann, Patrick Sapel, Ina Berenice Fink, Setz Schmitz, Christian Hopmann, Matthias Jarke, Günther Schuh, Klaus Wehrle, and Martin Henze. Designing Secure and Privacy-Preserving Information Systems for Industry Benchmarking. `https://github.com/COMSYS/industry-benchmarking`, 2023.

[SrcC24a] Jan Pennekamp, Fritz Alder, Lennart Bader, Gianluca Scopelliti, Klaus Wehrle, and Jan Tobias Mühlberg. Securing Sensing in Supply Chains: Opportunities, Building Blocks, and Designs. `https://github.com/COMSYS/secure-sensing`, 2024.

[SrcC24b] Lennart Bader, Jan Pennekamp, Roman Matzutt, and Klaus Wehrle. PrivAcciChain: Blockchain-Based Privacy Preservation for Supply Chains Supporting Lightweight Multi-Hop Information Accountability. `https://github.com/COMSYS/PrivAccIChain`, 2024.

[SS94] Ravi S. Sandhu and Pierangela Samarati. Access Control: Principle and Practice. *IEEE Communications Magazine*, 32(9):40–48, 1994.

[SS02] Togar M. Simatupang and Ramaswami Sridharan. The Collaborative Supply Chain. *The International Journal of Logistics Management*, 13(1):15–30, 2002.

[SSH+18] Emiliano Sisinni, Abusayeed Saifullah, Song Han, Ulf Jennehag, and Mikael Gidlund. Industrial Internet of Things: Challenges, Opportunities, and Directions. *IEEE Transactions on Industrial Informatics*, 14(11):4724– 4734, 2018.

[SSK+13] Axel Schroepfer, Andreas Schaad, Florian Kerschbaum, Heiko Boehm, and Joerg Jooss. Secure Benchmarking in the Cloud. In *Proceedings of the 18th ACM Symposium on Access Control Models and Technologies (SACMAT '13)*, pages 197–200. ACM, 2013.

[SSK14] Kamil Salikhov, Gustavo Sacomoto, and Gregory Kucherov. Using cascading Bloom filters to improve the memory usage for de Brujin graphs. *Algorithms for Molecular Biology*, 9(1), 2014.

[SSS17] Lakshmi Siva Sankar, M. Sindhu, and M. Sethumadhavan. Survey of consensus protocols on blockchain applications. In *Proceedings of the 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS '17)*. IEEE, 2017.

[SSS23] Maria Spiß, Tobias Schröer, and Günther Schuh. Resilience Configurator for Procurement. In *Proceedings of the IFIP International Conference on Advances in Production Management Systems (APMS '23)*, volume 692, pages 699–713. Springer, 2023.

[Sta96] Markus Stadler. Publicly Verifiable Secret Sharing. In *Proceedings of the International Conference on the Theory and Application of Cryptographic Techniques (EUROCRYPT '96)*, volume 1070, pages 190–199. Springer, 1996.

[Sto97] Donald E. Stokes. *Pasteur's Quadrant: Basic Science and Technological Innovation*. Brookings Institution Press, 1st edition, 1997.

[SW05] Ralf Steinmetz and Klaus Wehrle. *Peer-to-Peer Systems and Applications*. Springer, 1st edition, 2005.

[Swe02] Latanya Sweeney. k-anonymity: a model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(5):557–570, 2002.

[SWGW20] Martin Serror, Eric Wagner, René Glebke, and Klaus Wehrle. QWIN: Facilitating QoS in Wireless Industrial Networks Through Cooperation. In *Proceedings of the 19th IFIP Networking 2020 Conference (NETWORKING '20)*, pages 386–394. IEEE, 2020.

[SWP00] Dawn Xiaoding Song, David Wagner, and Adrian Perrig. Practical Techniques For Searches On Encrypted Data. In *Proceedings of the 2000 IEEE Symposium on Security and Privacy (SP '00)*, pages 44–55. IEEE, 2000.

[SWW15] Ahmad-Reza Sadeghi, Christian Wachsmann, and Michael Waidner. Security and Privacy Challenges in Industrial Internet of Things. In *Proceedings of the 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC '15)*. ACM, 2015.

[TBP23] Daniel Trauth, Thomas Bergs, and Wolfgang Prinz. *The Monetization of Technical Data: Innovations from Industry and Research*. Springer, 1st edition, 2023.

[TDDFD20] Koen Tange, Michele De Donno, Xenofon Fafoutis, and Nicola Dragoni. A Systematic Survey of Industrial Internet of Things Security: Requirements and Fog Computing Opportunities. *IEEE Communications Surveys & Tutorials*, 22(4):2489–2520, 2020.

[TDKCK22] Hoa Tran-Dang, Nicolas Krommenacker, Patrick Charpentier, and Dong-Seong Kim. The Internet of Things for Logistics: Perspectives, Application Review, and Challenges. *IETE Technical Review*, 39(1):93–121, 2022.

[Tei01] Eric Teicholz. *Facility Design and Management Handbook*. McGraw-Hill, 1st edition, 2001.

[TGH+18] Hasan Tercan, Alexandro Guajardo, Julian Heinisch, Thomas Thiele, Christian Hopmann, and Tobias Meisen. Transfer-Learning: Bridging the Gap between Real and Simulation Data for Machine Learning in Injection Molding. *Procedia CIRP*, 72:185–190, 2018.

[The19] The Linux Foundation. Confidential Computing Consortium – Linux Foundation Project. `https://confidentialcomputing.io/`, 2019.

[The20] The Apache Software Foundation. Apache HTTP Server Version 2.4 Documentation. `https://httpd.apache.org/docs/2.4/en/`, 2020.

[The23] The Economist. The world's most valuable resource is no longer oil, but data. `https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data`, 2017 (accessed April 4, 2023).

[TJA10] Hassan Takabi, James B. D. Joshi, and Gail-Joon Ahn. Security and Privacy Challenges in Cloud Computing Environments. *IEEE Security & Privacy*, 8(6):24–31, 2010.

[TLK+18] Muoi Tran, Loi Luu, Min Suk Kang, Iddo Bentov, and Prateek Saxena. Obscuro: A Bitcoin Mixer using Trusted Execution Environments. In *Proceedings of the 34th Annual Computer Security Applications Conference (ACSAC '18)*, pages 692–701. ACM, 2018.

[TZC87] Frances Gaither Tucker, Seymour M. Zivan, and Robert C. Camp. How to Measure Yourself Against the Best. *Harvard Business Review*, 65(1):8–10, 1987.

[UMNE+16] Nils Ulltveit-Moe, Henrik Nergaard, László Erdödi, Terje Gjøsæter, Erland Kolstad, and Pål Berg. Secure Information Sharing in an Industrial Internet of Things. arXiv:1601.04301, 2016.

[uRYS+19] Muhammad Habib ur Rehman, Ibrar Yaqoob, Khaled Salah, Muhammad Imran, Prem Prakash Jayaraman, and Charith Perera. The role of big data analytics in industrial Internet of Things. *Future Generation Computer Systems*, 99:247–259, 2019.

[van16] Wil M. P. van der Aalst. *Process Mining: Data Science in Action.* Springer, 2nd edition, 2016.

[VBMP17] Jo Van Bulck, Jan Tobias Mühlberg, and Frank Piessens. VulCAN: Efficient Component Authentication and Software Isolation for Automotive Control Networks. In *Proceedings of the 33rd Annual Computer Security Applications Conference (ACSAC '17)*, pages 225–237. ACM, 2017.

[VBOM+19] Jo Van Bulck, David Oswald, Eduard Marin, Abdulla Aldoseri, Flavio D. Garcia, and Frank Piessens. A Tale of Two Worlds: Assessing the Vulnerability of Enclave Shielding Runtimes. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19)*, pages 1741—1758. ACM, 2019.

[vdA19] Wil MP van der Aalst. Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data. In *Proceedings of the 17th International Conference on Software Engineering and Formal Methods (SEFM '19)*, volume 11724, pages 3–25. Springer, 2019.

[vdA21] Wil M. P. van der Aalst. Federated Process Mining: Exploiting Event Data Across Organizational Boundaries. In *Proceedings of the 2021 IEEE International Conference on Smart Data Services (SMDS '21)*. IEEE, 2021.

[vdAJKQ23] Wil M. P. van der Aalst, Matthias Jarke, István Koren, and Christoph Quix. Digital Shadows: Infrastructuring the Internet of Production. In *Internet of Production: Fundamentals, Applications and Proceedings*, pages 17–33. Springer, 2023.

[VDGHV10] Marten Van Dijk, Craig Gentry, Shai Halevi, and Vinod Vaikuntanathan. Fully Homomorphic Encryption over the Integers. In *Proceedings of the 29th Annual International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT '10)*, volume 6110, pages 24–43. Springer, 2010.

[vdKAB+18] Erik van der Kouwe, Dennis Andriesse, Herbert Bos, Cristiano Giuffrida, and Gernot Heiser. Benchmarking Crimes: An Emerging Threat in Systems Security. arXiv:1801.02381, 2018.

[VDM15] VDMA e.V. (Mechanical Engineering Industry Association). The VDMA – VDMA. `https://www.vdma.org/en/`, 2015.

[vGPZ22] Daniel van Geerenstein, Holger Paul, and Steffen Zimmermann. VDMA Product Piracy Study 2022. Yellow paper, VDMA, 2022.

[VJH21] Alexander Viand, Patrick Jattke, and Anwar Hithnawi. SoK: Fully Homomorphic Encryption Compilers. In *Proceedings of the 2021 IEEE Symposium on Security and Privacy (SP '21)*, pages 1092–1108. IEEE, 2021.

[VKH23] Alexander Viand, Christian Knabenhans, and Anwar Hithnawi. Verifiable Fully Homomorphic Encryption. arXiv:2301.07041, 2023.

[VKW+16] Diana M. Segura Velandia, Navjot Kaur, William G. Whittow, Paul P. Conway, and Andrew A. West. Towards industrial internet of things: Crankshaft monitoring, traceability and tracking using RFID. *Robotics and Computer-Integrated Manufacturing*, 41:66–77, 2016.

[Vla22] Eduard Vlad. Applying Trusted Execution for Privacy-Preserving Company Benchmarking. Bachelor's thesis, RWTH Aachen University, 2022.

[VMPL21] Bastian Verhaelen, F. Mayer, Sina Peukert, and Gisela Lanza. A comprehensive KPI network for the performance measurement and management in global production networks. *Production Engineering*, pages 635–650, 2021.

[VTF+18] Herman Voigts, Daniel Trauth, Andreas Feuerhack, Patrick Mattfeld, and Fritz Klocke. Dependencies of the die-roll height during fine blanking of case hardening steel 16MnCr5 without V-ring using a nesting strategy. *The International Journal of Advanced Manufacturing Technology*, 95(5):3083–3091, 2018.

[Wag20] Simon Wagner. Privacy-Preserving Company Benchmarking. Bachelor's thesis, RWTH Aachen University, 2020.

[WDA+16] Mark D. Wilkinson, Michel Dumontier, I. Jsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 2016.

[WG18] Karl Wüst and Arthur Gervais. Do you need a Blockchain? In *Proceedings of the 2018 Crypto Valley Conference on Blockchain Technology (CVCBT '18)*, pages 45–54. IEEE, 2018.

[Wit17] Krzysztof Witkowski. Internet of Things, Big Data, Industry 4.0 – Innovative Solutions in Logistics and Supply Chains Management. *Procedia Engineering*, 182:763–769, 2017.

[WKW16] Karl Weiss, Taghi M. Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big Data*, 3(1):9, 2016.

[WKW+23] Konrad Wolsing, Dominik Kus, Eric Wagner, Jan Pennekamp, Klaus Wehrle, and Martin Henze. One IDS is not Enough! Exploring Ensemble Learning for Industrial Intrusion Detection. In *Proceedings of the 28th European Symposium on Research in Computer Security (ESORICS '23)*, volume 14345, pages 102–122. Springer, 2023.

[WLC14] David Wood, Markus Lanthaler, and Richard Cyganiak. RDF 1.1 Concepts and Abstract Syntax. W3C Rec., 2014.

[WM11] Michael E. Whitman and Herbert J. Mattord. *Principles of Information Security*. Course Technology Press, 4th edition, 2011.

[WMP+22] Eric Wagner, Roman Matzutt, Jan Pennekamp, Lennart Bader, Irakli Bajelidze, Klaus Wehrle, and Martin Henze. Scalable and Privacy-Focused Company-Centric Supply Chain Management. In *Proceedings of the 2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC '22)*. IEEE, 2022.

[WNYD09] Dong-Liang Wu, Wing W. Y. Ng, Daniel S. Yeung, and Hai-Lan Ding. A brief survey on current RFID applications. In *Proceedings of the 2009 International Conference on Machine Learning and Cybernetics (ICMLC '09)*, volume 4, pages 2330–2335. IEEE, 2009.

[Woo14] Gavin Wood. Ethereum: A Secure Decentralised Generalised Transaction Ledger. Yellow paper, Ethereum, 2014.

[WR10] Katinka Wolter and Philipp Reinecke. Performance and Security Tradeoff. *Proceedings of the 10th International School on Formal Methods for the Design of Computer, Communication and Software Systems (SFM '10)*, 6154:135–167, 2010.

[WR17] Ulrich Weigel and Marco Ruecker. *The Strategic Procurement Practice Guide*. Springer, 2017.

[WS23] Rene Wagner and Nadine Schimroszik. Ukraine conflict adds to European supply chain snags. `https://www.reuters.com/markets/europe/ukraine-conflict-adds-european-supply-chain-snags-2022-03-09/`, 2022 (accessed April 4, 2023).

[WSJ17] Martin Wollschlaeger, Thilo Sauter, and Jürgen Jasperneite. The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0. *IEEE Industrial Electronics Magazine*, 11(1):17–27, 2017.

[WTvS+22] Konrad Wolsing, Lea Thiemt, Christian van Sloun, Eric Wagner, Klaus Wehrle, and Martin Henze. Can Industrial Intrusion Detection Be SIMPLE? In *Proceedings of the 27th European Symposium on Research in Computer Security (ESORICS '22)*, volume 13556, pages 574–594. Springer, 2022.

[WVK18] Martin Westerkamp, Friedhelm Victor, and Axel Küpper. Blockchain-Based Supply Chain Traceability: Token Recipes Model Manufacturing Processes. In *Proceedings of the 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pages 1595–1602. IEEE, 2018. Proceedings of the 2018 IEEE International Conference on Blockchain (Blockchain '18).

[WWSH22] Konrad Wolsing, Eric Wagner, Antoine Saillard, and Martin Henze. IPAL: Breaking up Silos of Protocol-dependent and Domain-specific Industrial Intrusion Detection Systems. In *Proceedings of the 25th International Symposium on Research in Attacks, Intrusions and Defenses (RAID '22)*, pages 510–525. ACM, 2022.

[XCK08] Yu Xia, Bintong Chen, and Panos Kouvelis. Market-Based Supply Chain Coordination by Matching Suppliers' Cost Structures with Buyers' Order Profiles. *Management Science*, 54(11):1861–1875, 2008.

[Xu17] Xun Xu. Machine Tool 4.0 for the new era of manufacturing. *The International Journal of Advanced Manufacturing Technology*, 92(5):1893–1900, 2017.

[XVHS20] Wanli Xue, Dinusha Vatsalan, Wen Hu, and Aruna Seneviratne. Sequence Data Matching and Beyond: New Privacy-Preserving Primitives Based on Bloom Filters. *IEEE Transactions on Information Forensics and Security*, 15:2973–2987, 2020.

[YMB+15] Fan Yang, Nelson Matthys, Rafael Bachiller, Sam Michiels, Wouter Joosen, and Danny Hughes. muPnP: Plug and Play Peripherals for the Internet of Things. In *Proceedings of the 10th European Conference on Computer Systems (EuroSys '15)*. ACM, 2015.

[YTR20] Jiho Yoon, Srinivas Talluri, and Claudia Rosales. Procurement decisions and information sharing under multi-tier disruption risk in a supply chain. *International Journal of Production Research*, 58(5):1362–1383, 2020.

[YXSW18] Chunyong Yin, Jinwen Xi, Ruxia Sun, and Jin Wang. Location Privacy Protection Based on Differential Privacy Strategy for Big Data in Industrial Internet of Things. *IEEE Transactions on Industrial Informatics*, 14(8):3628–3636, 2018.

[ZC11] Liang Zhou and Han-Chieh Chao. Multimedia Traffic Security Architecture for the Internet of Things. *IEEE Network*, 25(3):35–40, 2011.

[ZC20] Xu Zheng and Zhipeng Cai. Privacy-Preserved Data Sharing Towards Multiple Parties in Industrial IoTs. *IEEE Journal on Selected Areas in Communications*, 38(5):968–979, 2020.

[ZDJ+15] Yuchen Zhang, Wenrui Dai, Xiaoqian Jiang, Hongkai Xiong, and Shuang Wang. FORESEE: Fully Outsourced secuRe gEnome Study basEd on homomorphic Encryption. *BMC Medical Informatics and Decision Making*, 15 (Suppl 5), 2015.

[ZDX+21] Yinghui Zhang, Robert H. Deng, Shengmin Xu, Jianfei Sun, Qi Li, and Dong Zheng. Attribute-based Encryption for Cloud Computing Access Control: A Survey. *ACM Computing Surveys*, 53(4), 2021.

[ZHHW15] Jan Henrik Ziegeldorf, Martin Henze, René Hummen, and Klaus Wehrle. Comparison-based Privacy: Nudging Privacy in Social Media (Position Paper). In *Proceedings of the 10th International Workshop on Data Privacy Management (DPM '15)*, volume 9481, pages 226–234. Springer, 2015.

[Zil20] Noa Zilberman. An Artifact Evaluation of NDP. *ACM SIGCOMM Computer Communication Review*, 50(2):32–36, 2020.

[ZKS+18] Yuanyu Zhang, Shoji Kasahara, Yulong Shen, Xiaohong Jiang, and Jianxiong Wan. Smart Contract-Based Access Control for the Internet of Things. *IEEE Internet of Things Journal*, 6(2):1594–1605, 2018.

[ZMJ+19] Johannes Zrenner, Frederik Oliver Möller, Christian Jung, Andreas Eitel, and Boris Otto. Usage control architecture options for data sovereignty in business ecosystems. *Journal of Enterprise Information Management*, 32(3):477–495, 2019.

[ZMR18] Mahdi Zamani, Mahnush Movahedi, and Mariana Raykova. RapidChain: Scaling Blockchain via Full Sharding. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS '18)*, pages 931–948. ACM, 2018.

[ZNP15] Guy Zyskind, Oz Nathan, and Alex Sandy Pentland. Decentralizing Privacy: Using Blockchain to Protect Personal Data. In *Proceedings of the 2015 IEEE Symposium on Security and Privacy Workshops (SPW '15)*, pages 180–184. IEEE, 2015.

[ZPH+17] Jan Henrik Ziegeldorf, Jan Pennekamp, David Hellmanns, Felix Schwinger, Ike Kunze, Martin Henze, Jens Hiller, Roman Matzutt, and Klaus Wehrle. BLOOM: BLoom filter based Oblivious Outsourced Matchings. *BMC Medical Genomics*, 10 (Suppl 2):29–42, 2017.

[ZPMG19] Zach Zacharia, Michael Plasch, Usha Mohan, and Markus Gerschberger. The emerging role of coopetition within inter-firm relationships. *The International Journal of Logistics Management*, 30(2):414–437, 2019.

[ZRC+18] Muwei Zheng, Hannah Robbins, Zimo Chai, Prakash Thapa, and Tyler Moore. Cybersecurity Research Datasets: Taxonomy and Empirical Analysis. In *Proceedings of the 11th USENIX Workshop on Cyber Security Experimentation and Test (CSET '18)*. USENIX Association, 2018.

[ZWD+12] Yong Zeng, Lingyu Wang, Xiaoguang Deng, Xinlin Cao, and Nafisa Khundker. Secure collaboration in global design and supply chain environment: Problem analysis and literature review. *Computers in Industry*, 63(6):545–556, 2012.

[ZXD+18] Zibin Zheng, Shaoan Xie, Hong-Ning Dai, Xiangping Chen, and Huaimin Wang. Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services*, 14(4):352–375, 2018.

[ZZD18] Yinghui Zhang, Dong Zheng, and Robert H. Deng. Security and Privacy in Smart Health: Efficient Policy-Hiding Attribute-Based Access Control. *IEEE Internet of Things Journal*, 5(3):2130–2145, 2018.

[ZZZ19] Qide Zheng, Xincun Zhuang, and Zhen Zhao. State-of-the-art and future challenge in fine-blanking technology. *Production Engineering*, 13(1):61–70, 2019.