Shaping the Digital Transformation in Production: An Information- and Network-Centric Perspective

✉ Jan Pennekamp[1], Ike Kunze[1], Liam Tirpitz[2], Benedikt Bode[1], Constantin Sander[1], Sandra Geisler[2], and Klaus Wehrle[1]

[1] *Communication and Distributed Systems*, RWTH Aachen University, Aachen, Germany `{lastname}@comsys.rwth-aachen.de`
[2] *Data Stream Management and Analysis*, RWTH Aachen University, Aachen, Germany `{lastname}@cs.rwth-aachen.de`

Abstract. The Internet of Production exemplifies the ongoing digital transformation in production by holistically approaching the creation of data-to-knowledge pipelines—from the initial collaborative collection of data from various, possibly competing entities up to the final knowledge generation and exploitation. Primary envisioned benefits include process optimization and resilience improvement to advance sustainability—addressing today's most pressing challenge. In this paper, we outline the current state of the digital transformation, highlighting research achievements across various layers of interest, namely application, data modeling, and the underlying infrastructure. Based on this assessment, we identify critical information- and network-centric questions that help guide the industrial domain to an even more collaborative and sustainable future.

Keywords: secure industrial collaboration · decentralized information dissemination · ad-hoc compute orchestration · process mining

1 Introduction

The digital transformation in production is an ongoing process, as exemplified by the Internet of Production (IoP) [8,38] and other large-scale initiatives. Aiming to improve data exploitation and the transition from data to knowledge, the IoP has made focused advances that enable sharing and jointly leveraging data and knowledge, partially even on a global scale and incorporating multiple, potentially competing stakeholders [5,29,35]. These advancements provide a solid foundation for evolving the industrial domain. However, corresponding developments face new fundamental research challenges that greatly impact the ongoing transformation.

Aiming to establish a knowledge-fused distributed business ecosystem that considers various forms of sustainability, several dimensions become important. First, while our previous work has demonstrated that secure industrial collaboration is possible, we initially only targeted designing use case-specific solutions, leaving universally applicable concepts still missing. Second, the IoP primarily

---

Dedicated to *Wil van der Aalst*.

This version of the contribution has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: `https://doi.org/10.1007/978-3-032-17618-9_39`. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use `https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms`.

focuses on creating well-defined collaboration environments and sourcing information from these interactions, limiting responsiveness to world-wide events and overall flexibility. Third, current initiatives tend to concentrate on fundamentally building pipelines without adequately addressing their sustainability; holistic solutions could utilize available resources more efficiently. Given this context, we identify three information- and network-centric questions:

- ► *How to conceptualize, evolve, and standardize secure industrial collaboration?*
- ► *Which approaches promise a successful signaling mechanism for global use?*
- ► *What is needed to enable and optimize a novel federated cloud continuum?*

Tackling these currently-overlooked research questions is crucial for reliably, safely, securely, and sustainably providing the right information with appropriate expressiveness to the right party at the right time—aspects that are essential for higher-layer analyses, decision-making, and operation in the context of production.

In this paper, we derive a corresponding research agenda that strives for a truly interconnected production landscape while simultaneously incorporating sustainability considerations in knowledge-fused distributed business ecosystems. For this, we first outline the status quo of the digital transformation in production, which has evolved considerably due to the emergence of approximated close-to-edge control loops, Models-in-the-Middle, and secure industrial collaboration [5,29,35,58]. Specifically, we recap key conceptual considerations of the IoP in Sect. 2 and distill its research impact concerning the realization of a World-Wide Lab from an information- and network-centric perspective with a focus on the data-to-knowledge approach in Sect. 3. In this context, we also highlight corresponding implications for higher-layer applications and decision-making. This summary serves as a point of origin for deriving a concrete agenda that further evolves the production landscape (Sect. 4), i.e., continues the digital transformation. We focus our presentation on an information- and network-centric viewpoint to create the foundation for higher-layer data-to-knowledge pipelines and improved decision-making, as we outline in Sect. 5. Finally, Sect. 6 concludes this paper.

2 Premise and Concepts of the Internet of Production

The Internet of Production (IoP) pursues the vision of setting up world-wide data-driven, cross-domain, and interorganizational collaboration by providing semantically adequate and context-aware data to relevant stakeholders in real time and at a reasonable level of granularity [38]. These dataflows are then embedded and orchestrated with the World-Wide Lab (WWL) [8], introducing significant benefits in terms of costs, productivity, flexibility, innovation, and product quality, among others [33], to participating organizations and society alike. Brauner et al. [8] classify corresponding developments for transforming production into four layers, which we visualize in Fig. 1.

In essence, data processed in the lowest layer is turned into information, which is then translated into knowledge, ready for communication to humans and organizations. This processing joins data, information, and knowledge from

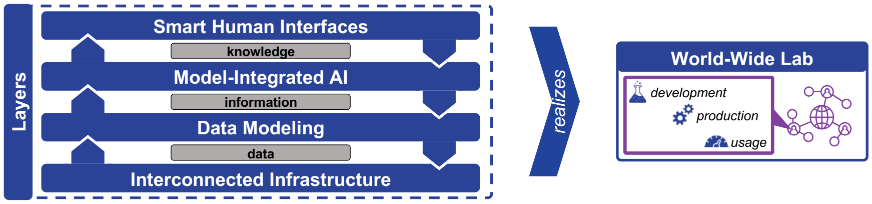

Fig. 1. In 2022, Brauner et al. [8] defined four research layers that shape the digital transformation in production, as the IoP exemplifies, and collectively realize the WWL.

previously organizationally-isolated phases, i.e., development, production, and usage, and builds on data abstraction, aggregation, and refinement [8]. The WWL complements these layers by facilitating seamless integration of data, information, and knowledge across stakeholders, while the layers contribute the following functionality:

- **Smart Human Interfaces:** Evolved interfaces enable both experts and non-experts to interact adaptively with generated knowledge and their environments. They are specifically tailored to tasks, contexts, and users, enhancing engagement and efficiency, for instance, by applying multi-faceted visual process analytics and multi-layer frameworks such as Tiramisù [2, 3].
- **Model-Integrated Artificial Intelligence (AI):** Using abstraction and aggregation, this layer proposes novel (foundational) models that offer tailored insights into production environments and their dependencies based on data, metadata, and (structured) information representations. It provides the knowledge basis for the smart human interfaces above. These interfaces can operationalize model outputs via layered, on-demand visual dimensions [2, 3].
- **Data Modeling:** This layer features abstraction and aggregation of data, which is pivotal for transforming raw data into information and actionable knowledge, including the abstraction of raw IoT sensor streams into discrete events for process mining [9]. As such, it fuels the AI models situated on top.
- **Interconnected Infrastructure:** The underlying infrastructure serves as the backbone of the IoP and WWL by ensuring reliable data sensing, retrieval, sharing, and processing, thus providing the basis for any data modeling. Systematic approaches to transform sensor data into event logs are particularly critical for IoT-rich environments [9], as prevalent in the interconnected IoP.

With organizational changes [27, 40] that enable the underlying exchange of information and data, these layers create the basis for collaboration-based value creation and facilitate the aforementioned cost, flexibility, and quality benefits.

Here, as a central layer, *Model-Integrated AI* is crucial for achieving the objectives because it is responsible for outputting findings for improved decision-making. Apart from machine learning, leveraging methodologies known from process mining promises to bridge the gap between model-based process analyses and data-oriented analyses [54]. Respective research activities on data-to-knowledge pipelines [15] are particularly valuable for complex environments with high concurrency. An early idea in this context covers the concept of process cubes [53], where events and process models are organized using different dimensions to discover, analyze, and improve (business) processes based on event data. Complementary approaches on visual process analytics explicitly address multi-faceted event data and interactive analytical abstraction [12]. While the *Smart Human Interfaces* layer utilizes this information to communicate derived knowledge to the stakeholders, the *Data Modeling* layer is particularly of interest for efficiently and reliably providing models to the environment that bridge data to knowledge and vice versa, i.e., connecting the surrounding layers.

By design, said layer builds upon the availability of information and, thus, initial data for its successful operation. Consequently, the *Interconnected Infrastructure* layer is essential for creating, providing, and maintaining added value in higher-level layers and applications like process mining. Specifically, the availability of rich data from diverse and potentially distributed yet broadly dependable data sources enables global optimization of processes by converting data into knowledge. Reviews of process mining on sensor data detail these pipelines and common pitfalls [9]. Beyond standardization of (secured) dataflows, we see a great demand for mature signaling and optimization approaches within this layer, as we substantiate and detail in Sect. 4. Thus, we emphasize the need to focus on this information- and network-centric perspective to establish a robust foundation for reliably collecting, handling, and sharing data safely and securely. Contrary actions will otherwise impair achieving the full potential of the IoP.

Having these overarching conceptual considerations of the IoP in mind, we next examine how its progress has shaped the digital transformation in production since the publication by Brauner et al. [8] in 2022.

3 Status Quo: Research Contributions and Achievements

Research within the IoP has made significant strides in advancing the understanding and application of digital shadows [4, 26], the WWL, and related concepts. In this section, we highlight key research contributions to the fundamental conceptualization and realization of the IoP and leave aside research on individual production processes or similar topics with a narrower scope. Specifically, we focus on innovations in three key domains that advance the overarching objective of the IoP of turning data into knowledge [58]: (a) data-to-knowledge pipelines, (b) interoperable data models, and (c) data communication and dataflows.

Data-to-Knowledge Pipelines [5]. The development of robust data-to-knowledge pipelines is a cornerstone component of the IoP. At their core, these pipelines rely on integrating real-time manufacturing data, simulations, and mathematical models, which yield the basis for communication and data sharing within the WWL. Specifically, autonomous AI agents can query the WWL for machine-level and process-level information, subsequently transforming the standalone raw (production) data into actionable insights. The focus is on creating systematic approaches that leverage machine learning and model-based AI methods to enhance decision-making processes for operators and managers alike, ultimately leading to the realization of validated self-adaptive production systems. For instance, utilizing these pipelines for predictive maintenance can significantly reduce downtime by enabling proactive adjustments based on real-time analytics. At the same time, these pipelines have the ability to analyze even complex process chains using process mining techniques. They can further complement data-driven approaches with manufacturing-specific structural information to generate comprehensive views of shop-floor processes. With this domain-oriented application, the IoP can focus on time-dependent metrics over traditional control-flow perspectives, hence, creating the foundation for instantiating object-centric process mining (OCPM) [55, 57] in production.

Interoperable Data Models [29]. Another critical area of advancement and driver of a successful IoP is the establishment of interoperable data models that promote seamless communication across diverse systems within the ecosystem. Specifically, the IoP focuses on conceptualizing an adequate set of data modeling techniques and transformations that link to metamodels to appropriately capture their semantics and cover the entire product life cycle. These interoperable models not only support cross-domain collaboration but simultaneously enhance the reusability of existing knowledge across different applications in manufacturing processes. The such-generated foundation for data modeling is essential for enabling higher-layer activities, including process mining methods such as process discovery and prediction, but also for ensuring interoperability of data, information, and knowledge. Key for the successful interfacing between the different models and the higher-layer applications is the Model(s)-in-the-Middle concept [21, 49], which bridges semantic gaps between stakeholders and ultimately fosters a more cohesive ecosystem where insights are shared efficiently between organizations, enabling various stakeholders to interpret data consistently.

Advances to Data Communication and Dataflows [35]. Third, providing the underlying infrastructure for enabling data models and the WWL, prior work has also focused on improving the fundamental dataflows and data processing capabilities on different levels. For instance, researchers have studied and analyzed dataflows in the IoP [39] and within supply chains [34, 40]. Secure industrial collaborations (SICs) and deployable information security are prime examples of such new forms of industrial data sharing [33]. Even though several designs were proposed, including an approach [41] that can collect the confidentiality requirements of the involved stakeholders in a structured manner, additional dimensions, including organizational security as well as legal and economic aspects, must still be considered when fully maturing the concept of SICs [33]. On a lower level, research has studied how data stream processing can be orchestrated in a scalable manner, for example, proposing a distributed, locality-aware data stream processing architecture that provides a framework for flexibly distributing and scaling compute tasks [47]. At the same time, work on in-network computing (INC) has opened up networking devices as new resources for executing compute tasks, e.g., showing that bandwidths can be reduced by leveraging process semantics [52] and response times can be shortened by transforming data [22] or analyzing the process state [23] in the network. These contributions have also helped shape the overall research agenda for INC beyond industrial scenarios [24].

Main IoP Achievements. Overall, the pointed out IoP achievements encompass a holistic approach to transforming traditional manufacturing paradigms and production into agile, interconnected ecosystems capable of adapting swiftly to dynamically-changing demands. Specifically, the focus was on mitigating the trade-off between agility, cost efficiency, and quality, and enabling a new level of cross-domain collaboration in production. By combining process mining with simulation and machine learning at both event and aggregate levels, novel solutions allow for the proactive anticipation of production issues, improving operational planning and control [43]. Similarly, the Model(s)-in-the-Middle concept standardizes data-sharing structures across domain silos and has proven to be a critical step in enhancing data exchange mechanisms and facilitating interoperability [21, 49], and also enabling large-scale data interoperability that builds upon the FAIR principles [13]. Finally, first communication system and information security realizations [33] as well as network security considerations [11] establish crucial building blocks for the establishment of a WWL that enables communication among numerous, heterogeneous stakeholders across the world.

Important Areas for Improvement. Despite these promising and broad advances, three relevant capabilities remain underdeveloped and limit the scalable, trustworthy realization of the WWL: (i) *Standardized Large-Scale Secure Industrial Collaborations:* Prior work has proposed several designs for SIC, but many solutions have been tailored to specific use cases, limiting the widespread adoption of the concept as such; consequently, universally-applicable, interoperable frameworks and standards for modeling, establishing, and operating SICs have yet to be developed and adopted. (ii) *Decentralized Information and Event Dissemination:* While research has described dataflows in the IoP and supply chains, today's established messaging channels lack support for automatically and securely setting up cross-domain channels, thereby constraining the utilization of valuable data and information. (iii) *Ad-hoc Compute Orchestration within the Cloud Continuum:* Locality-aware stream processing and INC provide building blocks for ad-hoc compute orchestration, but policy-aware placement, migration, and monitoring across the edge-network-cloud continuum and across domains are not yet fully explored. By advancing the WWL, these areas for improvement promise to significantly realize the vision of the IoP in practice. In the next section, we thus follow up on them to outline a stringent research agenda.

4 An Information- and Network-Centric Research Agenda for Shaping the Digital Transformation in Production

After establishing the status quo of the digital transformation in production, we now dive into its consecutive research agenda. Specifically, we still focus on information- and network-centric aspects, i.e., the lowest layer of the transformation (cf. Fig. 1)—the interconnected infrastructure—but now turn our attention toward future lines of research. Building upon the first successes related to SICs, we discuss the next steps in maturing and standardizing this technology in Sect. 4.1. In an effort to further evolve these collaborations, we present our vision for a global signaling approach in Sect. 4.2 that eases and maintains information flows in the WWL. We then examine the challenges of holistically co-optimizing computation, its placement across compute nodes, the corresponding data transmissions, and the concurrent selection of algorithms in Sect. 4.3. Finally, in Sect. 4.4, we summarize our information- and network-centric research directions and highlight shared sustainability considerations and goals.

4.1 Standardized Large-Scale Secure Industrial Collaborations

The industrial landscape, participating organizations, and their security and privacy requirements are very diverse. As such, different industrial scenarios and information flows usually require individual solutions as we also showed with our work on use case-specific SICs [33]. In particular, we relied on different technical building blocks [39], applying both software- and hardware-based ones, to provide organizations with reliable (security) guarantees. At the same time, research best practices [20, 36] encourage the reuse of previous work, and we showed that reuse across domains or use cases may still be possible, potentially with minor adaptations (e.g., exchanging production process parameters or milling tool specifications [37]). These considerations motivated us to develop ConfMod [41], a middleware that simplifies and standardizes a fine-grained modeling of confidentiality requirements.

Based on these observations, the next step of the transformation is to standardize SICs. To this end, one way may be to create a toolbox that captures different (conceptual) approaches for collaborating securely among multiple, potentially even mutually distrusting, organizations in a reliable and interoperable manner. This toolbox can effectively hold approaches for various use cases and settings, showcasing the breadth of SICs. By applying ConfMod, stakeholders can then model their confidentiality requirements in a standardized and reusable representation to identify other use cases with similar constraints and eventually apply (and adapt if needed) a well-known approach for their respective deployment. In the future, ConfMod could (a) further contribute to ensuring that derived deployments comply with competition laws and data protection regulations, and even (b) derive accurate, legally-binding cooperation agreements that additionally safeguard SICs. Both measures are likely to accelerate the real-world use of SICs.

Certainly, this SIC toolbox needs to evolve over time to account for the most recent developments in terms of scale (with a growing WWL, approaches need to scale accordingly [18, 48]), technology (large-scale federated approaches are only beginning to emerge [25, 56]), and level of automation (while early collaborations focused on comparing information, more sophisticated ones will autonomously control processes [35]). For instance, we expect an evolution from bilateral SICs, whose data is processed manually, to massively-federated, potentially cloud-based SICs that exploit information more broadly—sourcing the entire ecosystem—in an automated manner. Given the lack of dependable reliability guarantees so far, organizations are still concerned with utilizing information from third parties without prior validation [35]. More advanced SICs may alleviate these concerns and thus have the potential of increasing the share of autonomously utilizing third-party-sourced information. We already see first attempts at utilizing and advancing federated machine learning [10,19] and federated process mining [16,44].

Despite the shift toward large-scale SICs, considering the resource footprint of collaborating (securely) is crucial. In particular, the security and communication overhead is immense, such that stakeholders should carefully weigh whether the additional resource usage is beneficial on a global scale. This question thus calls for benchmarking criteria that capture this trade-off timely and accurately.

4.2 Global Signaling and Decentralized Information Dissemination

While the envisioned SIC toolbox will ease and secure collaborations, getting notified of and finding relevant information is increasingly difficult due to the exponential growth of global data. Additionally, increasing interdependencies resulting from globalization can easily lead to ripple effects felt around the world, as was evident when the Suez Canal was blocked by a ship in 2021 [31]. To realize a truly global, sustainable, and adaptable WWL, efficient information exchange is crucial to meet the growing demand for meaningful information flows. This demand extends beyond merely establishing the flows but requires sharing signals and information within rapidly changing networks of entities with increasing scale and a high need for reliability and federation. Given these complexities, we argue that current approaches, such as popular message brokers like Kafka or basic data-sharing platforms, are inadequate: they lack semantic interoperability, scale poorly across sovereign domains, and provide limited support for federation. New methods are therefore required to enable a truly global, yet sovereign and adaptable, exploitation of knowledge through massively networked information.

We propose to address the identified need via decentralized information and event dissemination that can globally connect "data silos" and enable new information flows, even across domains. In particular, we envision a meaningful interweaving of information from diverse sources and stakeholders to ensure all entities receive appropriate and up-to-date information at the right time. Central to this concept is an embedded, decentralized signaling system inspired by the human nervous system, offering two complementary modes of operation: push (afferent) and pull (efferent). Unlike conventional message brokers, where producers push opaque messages into a centralized system, our approach enables semantically aware dissemination across sovereign domains. Information providers can push updates into the system, which then dynamically routes them to relevant recipients while accounting for transitive dependencies and potential ripple effects. At the same time, entities are able to pull relevant information as needed, ensuring both adaptability and sovereignty in global-scale information flows.

From a technical perspective, our dissemination framework relies on key conceptual ideas of information-centric networking (ICN) [1], an approach that fundamentally shifts networking to an information-centered view [30], enabling hosts to query specific information from the network without needing to know where exactly the information resides. Similarly, peer-to-peer-based approaches that utilize distributed hash tables, e.g., Chord [46], provide decentralized and scalable structures that help in realizing the system in a sovereign and fault-proof way. Beyond the fundamental signal and information exchange, we also consider additional important aspects, such as security and information protection. For example, we envision making use of attribute-based encryption (ABE) [7, 60] to safeguard information in flows and provide confidentiality. ABE allows for selectively granting access without the need to specify individual recipients (per message), instead relying on assigned attributes and carefully-crafted policies. Similarly, we intend to use such concepts to restrict the spread of information and to limit interests as much as possible. Overall, additional consideration of various network and security attacks (and how to securely embed security protocols) is needed to develop a system that protects against these threats.

4.3 Ad-hoc Compute Orchestration in Federated Cloud Continuum

The described signaling and information dissemination framework provides the foundation for a vision of dynamic (ad-hoc), large-scale coordination across the cloud–edge continuum. Rather than limiting its role to transporting raw data and notifications in a globally-distributed network of independent organizations, the framework can act as a control fabric that enables both information flows and computational workloads to be steered to the right place at the right time. By integrating computation into global data exchange, such an approach can reduce unnecessary data movement, preserve privacy, and optimize the use of limited compute and network resources distributed across the whole network [50].

In the cloud-edge continuum, workloads may be orchestrated vertically, across heterogeneous edge environments and centralized cloud infrastructures, or federated horizontally, across independent organizations and providers, forming a WWL. Deciding where and when to execute computation depends on multiple criteria: proximity to data sources, availability of compute and network resources, local energy conditions, and compliance with regulatory or organizational data sharing policies [51]. All these conditions can be signaled and acted upon through our event dissemination framework.

Globally-distributed workload placement offers several opportunities for advancing the federated cloud-edge continuum. For instance, computation can be shifted closer to data sources to uphold sovereignty and security, or redirected to regions with greener energy to minimize environmental impact [32], albeit at the cost of increased network utilization and reduced control. Considering the data-quality and quality-of-service requirements of applications, we can tune different dimensions, including resource utilization, accuracy and timeliness of results, and processing locality. For instance, through time-shifting execution, operators can align workloads with fluctuations in renewable energy availability. As a result, the carbon footprint can be reduced and different forms of efficiency, including energy efficiency and resource utilization, can be jointly optimized, although timeliness is lost [59]. Dynamic adjustments of the forwarding frequency and data resolution can be used to allow for more specific runtime trade-offs. Additionally, by reducing data volumes on the edge close to the data source, e.g.,

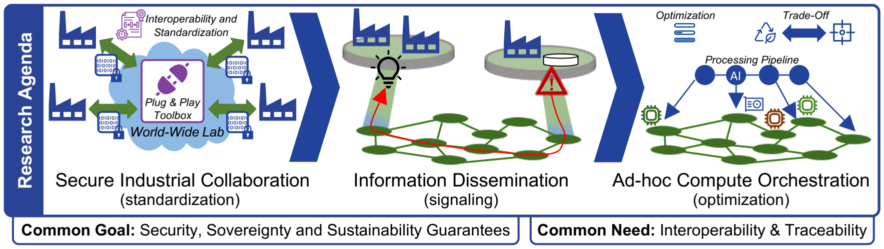

Fig. 2. The three research directions directly build on each other. Standardized collaboration methods enable dynamic and interoperable signaling across organizational boundaries, which is, in turn, the foundation for globally distributed processing pipelines.

by placing data stream synopses on the data path [42], the utilization of network links toward the cloud can be reduced and, depending on the context, latency can be improved. In this way, ad-hoc compute orchestration helps to dynamically balance performance, sustainability, and sovereignty.

At a global scale, our signaling framework enables independent systems to interoperate in flexible and spontaneous ways. Its capabilities allow participants to advertise resources, data availability, or sustainability attributes, while simultaneously discovering and engaging with partners that meet their own quality, trust, or policy objectives. As conditions evolve, e.g., when energy mixes change or new data sources appear, our framework supports reconfiguring pipelines and migrating workloads in an ad-hoc, on-demand fashion.

This vision positions our event dissemination framework as the backbone of a vertically and horizontally integrated cloud–edge ecosystem. Coupling dynamic data exchange with distributed workload orchestration enables global environments in which compute is actively steered toward globally optimal trade-offs across performance, sustainability, quality, and sovereignty.

4.4 Continued Transformation: Converging Global Collaboration, Signaling, and Optimization through Interoperable Data

In this section, we have highlighted a research agenda for the continued digital transformation in production based on three directions, as we also summarize in Fig. 2. An interoperable toolbox with adaptable methods for different settings can provide the foundation for large-scale, standardized secure industrial collaborations. These methods may be instantiated to move from rigid data exchange between selected organizations within the WWL to dynamic and distributed ad-hoc collaborations. The signaling of our event dissemination framework provides the necessary communication infrastructure to dynamically connect independent organizations at scale through interoperable information flows. Further, by building on cross-domain dataflows of our framework, we can globally distribute and orchestrate workloads in a federated cloud-edge continuum. Here, exchanging signals on available compute resources allows for ad-hoc optimization of global, collaborative processing pipelines.

Across all these efforts, two core themes emerge: First, sovereign infrastructures are a core pillar for globally-collaborative environments, safeguarding sensitive processes and information in computation and sharing, building trust between decentralized entities, and enabling secure and interoperable information flows and processing pipelines. Second, the ability to dynamically adapt setups based on given requirements is paramount for global optimization. A constant trade-off between sustainability, information availability, and sovereignty across independent systems is required to reduce energy consumption. For example, data quality can be reduced through local processing, which also reduces the communication frequency compared to centralized computation in the cloud. Locally computing approximated results or offloading compute to a third party can also optimize resource usage, including carbon emissions.

Through the previously described layers of industrial collaboration, event dissemination, and compute orchestration, we identified key common objectives.

Common Goal: *Providing security, sovereignty, and sustainability guarantees through trust, adaptability, and optimization in knowledge-fused distributed business ecosystems to improve production.*

Realizing interoperability and trust in these global ecosystems requires common data formats and traceability of physical and digital resources, as well as any applied transformations [6,14]. The Model(s)-in-the-Middle concept [21,49] can contribute to standardized data-sharing structures and therefore interoperable understanding of data across organizations. Building on top of these data models, data structures like the digital shadows [29] or digital product records [45] can contribute to the vision of a WWL in the future, by capturing traceable data through the entire life cycle of physical and digital products across organizations. In the future, such evolving records could log information on the initial manufacturing process and provenance of raw materials and their sustainability, but also continuously collect usage and reuse information, integrating data sources from various organizations along the supply and usage chain. Through standardized data models, the proposed event dissemination framework could base forwarding decisions on the content of product records and enable federated records by signaling updates and usage information across organizations, providing the foundation for higher-level applications built on a dynamic communication infrastructure and interoperable and traceable data records.

Common Need: *Interoperable data exchange formats and provenance information relating to the history and origin of physical and digital goods.*

5 Interplay with Higher-Layer Research Directions

By design (cf. Fig. 1), the research directions discussed in Sect. 4, concerning the information- and network-centric perspective, influence higher layers and vice versa. Specifically, we identify four main aspects of interplay.

First, information flows and relationships within the distributed business ecosystem provide diverse input for sophisticated process mining activities. For example, recent advances like OCPM [55,57] allow for conducting more accurate and holistic analyses concerning the design, maintenance, and evaluation of data-to-knowledge pipelines, as required for the IoP. Relatedly, SIC, our event dissemination framework, ad-hoc compute orchestration, and their associated information flows add new deployment areas for state-of-the-art approaches of the model-integrated AI layer, including federated process mining [56].

Second, compared to more traditional (process) optimization goals like productivity, product quality, or cost, the ongoing digital transformation in production adds another dimension—sustainability (cf. Sect. 4.4). Applying this dimension to process mining research could introduce a new optimization goal when looking at processes and event logs [17], thereby shifting the focus away from traditional control-flow perspectives. Made findings may even point out optimization potential for the lower layers as well. Ongoing activities already showcase the amplitude of this dimension in general [17]. Any efforts taken should be equally translatable to the lower layers, including the discussed future research directions.

Third, primarily concerning the data modeling layer, research should look into how (matured) concepts known from process mining can help track where information originates (and exploit this tracking), how it transforms, and where it flows, ensuring end-to-end provenance within the entire business ecosystem. In terms of our discussed research focus, corresponding activities are mainly related to SIC and data ownership questions and their legal implications. By addressing legal concerns and resolving uncertainties related to ownership and liability in distributed data-to-knowledge pipelines, advances in this direction may strengthen the perceived benefits of participating in the envisioned WWL.

Four, general feedback from and demands communicated by higher-layer applications within this distributed ecosystem in terms of deadlines, quality (e.g., resolution or approximation potential), and all sorts of information security and data sovereignty needs, including safeguarding data in transit and in use [28], introduce requirements for the lower layers. Accordingly, all concepts, designs, and implementations need to take them into account as part of their evolution.

To conclude, we observe direct interplay of the different layers that we introduced in Sect. 2 with varying implications for research and operation.

6 Conclusion

Based on the research roadmap by Brauner et al. [8], we set out to assess the state of digital transformation in production. In this context, the Internet of Production (IoP) is a prime example of a research initiative with strong ties to computer science research, ranging from improved process-mining applications over the conceptualization of Models-in-the-Middle to the application of novel processing paradigms like in-network computing. Even though our assessment of the status quo revealed significant advances (since 2022) in the ongoing transformation, we still identified three major challenges that result in a call for action on information- and network-centric research. First, we see the benefits of standardizing secure industrial collaborations to eventually deploy it on a large scale within the WWL. Second, we derive the need for a novel signaling mechanism that is able to reliably distribute information and knowledge among distributed stakeholders. Third, we call for an improved cloud continuum concept that optimizes the placement of computations. Thus, to continue the digital transformation in production, these research directions must consider the trade-off between the availability of data and sustainable operation while introducing diverse reliability, security, and sovereignty guarantees into knowledge-fused distributed business ecosystems.

Acknowledgments. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Sustainable Production – 390621612. We further thank all contributors to the *Internet of Production (IoP)* and *Internet of Sustainable Production (IoSP)* proposals, as well as all researchers who contribute(d) to their research objectives.

References

1. Ahlgren, B., Dannewitz, C., Imbrenda, C., Kutscher, D., Ohlman, B.: A Survey of Information-Centric Networking. IEEE Commun. Mag. **50**(7) (2012). `https://doi.org/10.1109/MCOM.2012.6231276`
2. Alman, A., Arleo, A., Beerepoot, I., Burattin, A., et al.: Tiramisù: A Recipe for Visual Sensemaking of Multi-faceted Process Information. In: Process Mining Workshops (2023). `https://doi.org/10.1007/978-3-031-56107-8_2`
3. Alman, A., Arleo, A., Beerepoot, I., Burattin, A., et al.: Tiramisù: making sense of multi-faceted process information through time and space. J. Intell. Inf. Syst. (2024). `https://doi.org/10.1007/s10844-024-00875-8`
4. Becker, F., Bibow, P., Dalibor, M., Gannouni, A., et al.: A Conceptual Model for Digital Shadows in Industry and Its Application. In: ER (2021). `https://doi.org/10.1007/978-3-030-89022-3_22`
5. Behery, M., Brauner, P., Zhou, H.A., Uysal, M.S., et al.: Actionable Artificial Intelligence for the Future of Production. In: Internet of Production: Fundamentals, Methods and Applications. Interdisciplinary Excellence Accelerator Series, Springer (2023). `https://doi.org/10.1007/978-3-031-44497-5{_}4`
6. Belova, A., Revenko, A., Gallina, V., Bachlechner, D., Sowe, S.K.: Bringing the Digital Product Passport to Life: Requirements Analysis for a Carbon Footprint Tracking System using Knowledge Graphs and Data Spaces. In: ACM SAC (2025). `https://doi.org/10.1145/3672608.3707971`
7. Bethencourt, J., Sahai, A., Waters, B.: Ciphertext-Policy Attribute-Based Encryption. In: IEEE SP (2007). `https://doi.org/10.1109/SP.2007.11`
8. Brauner, P., Dalibor, M., Jarke, M., Kunze, I., et al.: A Computer Science Perspective on Digital Transformation in Production. ACM Trans. Internet Things **3**(2) (2022). `https://doi.org/10.1145/3502265`
9. Brzychczy, E., Aleknonytė-Resch, M., Janssen, D., Koschmider, A.: Process mining on sensor data: a review of related works. Knowl. Inf. Syst. **67**(6) (2025). `https://doi.org/10.1007/s10115-024-02297-y`
10. Choudhary, S.K., Kar, A.K., Dwivedi, Y.K.: How does Federated Learning Impact Decision-Making in Firms: A Systematic Literature Review. Commun. Assoc. Inf. Syst. **54**(1) (2024). `https://doi.org/10.17705/1CAIS.05419`
11. Dahlmanns, M.: Identifying Security Issues in the Industrial Internet of Things. Ph.D. thesis, RWTH Aachen University (2025)

12. van den Elzen, S., Jans, M., Martin, N., Pieters, F., et al.: Towards Multi-Faceted Visual Process Analytics. Inf. Syst. **133** (2025). `https://doi.org/10.1016/j.is.2025.102560`
13. Gleim, L., Pennekamp, J., Liebenberg, M., Buchsbaum, M., et al.: FactDAG: Formalizing Data Interoperability in an Internet of Production. IEEE Internet Things J. **7**(4) (2020). `https://doi.org/10.1109/JIOT.2020.2966402`
14. Gleim, L., Tirpitz, L., Pennekamp, J., Decker, S.: Expressing FactDAG Provenance with PROV-O. In: MEPDaW (2020)
15. Gorißen, L., Schneider, J.N., Behery, M., Brauner, P., et al.: Demonstrating Data-to-Knowledge Pipelines for Connecting Production Sites in the World Wide Lab. arXiv:2412.12231 (2024). `https://doi.org/10.48550/arXiv.2412.12231`
16. Graves, N., Koren, I., Rafiei, M., van der Aalst, W.M.P.: From Identities to Quantities: Introducing Items and Decoupling Points to Object-Centric Process Mining. In: COMINDS (2023). `https://doi.org/10.1007/978-3-031-56107-8_35`
17. Graves, N., Koren, I., van der Aalst, W.M.P.: ReThink Your Processes! A Review of Process Mining for Sustainability. In: ICT4S (2023). `https://doi.org/10.1109/ICT4S58814.2023.00025`
18. Hazra, A., Adhikari, M., Amgoth, T., Srirama, S.N.: A Comprehensive Survey on Interoperability for IIoT: Taxonomy, Standards, and Future Directions. ACM Comput. Surv. **55**(1) (2021). `https://doi.org/10.1145/3485130`
19. Hirt, R., Kühl, N., Martin, D., Satzger, G.: Enabling inter-organizational analytics in business networks through meta machine learning. Inf. Technol. Manag. **26**(1) (2025). `https://doi.org/10.1007/s10799-023-00399-7`
20. Kim, S.Y., Tirpitz, L., Wagels, M., Arnold, B.T., et al.: Dataspaces for Collaborative Research. In: IEEE BigData (2025)
21. Koren, I., Jarke, M., Michael, J., Heithoff, M., et al.: Navigating the Data Model Divide in Smart Manufacturing: An Empirical Investigation for Enhanced AI Integration. In: EMMSAD (2024). `https://doi.org/10.1007/978-3-031-61007-3_21`
22. Kunze, I., Glebke, R., Scheiper, J., Bodenbenner, M., et al.: Investigating the Applicability of In-Network Computing to Industrial Scenarios. In: IEEE ICPS (2021). `https://doi.org/10.1109/ICPS49255.2021.9468247`
23. Kunze, I., Scheurenberg, D., Tirpitz, L., Geisler, S., Wehrle, K.: In-Situ Model Validation for Continuous Processes Using In-Network Computing. In: IEEE ICPS (2024). `https://doi.org/10.1109/ICPS59941.2024.10639999`
24. Kunze, I., Wehrle, K., Trossen, D., Montpetit, M.J., et al.: Use Cases for In-Network Computing. IRTF RFC 9817 (2025). `https://doi.org/10.17487/RFC9817`
25. Li, T., Sahu, A.K., Talwalkar, A., Smith, V.: Federated Learning: Challenges, Methods, and Future Directions. IEEE Signal Process. Mag. **37**(3) (2020). `https://doi.org/10.1109/MSP.2020.2975749`
26. Liebenberg, M., Jarke, M.: Information Systems Engineering with DigitalShadows: Concept and Case Studies. In: CAiSE (2020). `https://doi.org/10.1007/978-3-030-49435-3_5`
27. Lohmöller, J., Jeon, H., Hentschel, J., Wehrle, K., Pennekamp, J.: Between Promise and Practice: Challenges and Misperceptions of Applying Privacy Enhancing Technologies in Business Contexts. In: HICSS (2026)
28. Lohmöller, J., Pennekamp, J., Matzutt, R., Schneider, C.V., et al.: The Unresolved Need for Dependable Guarantees on Security, Sovereignty, and Trust in Data Ecosystems. Data Knowl. Eng. **151** (2024). `https://doi.org/10.1016/j.datak.2024.102301`

29. Michael, J., Koren, I., Dimitriadis, I., Fulterer, J., et al.: A Digital Shadow Reference Model for Worldwide Production Labs. In: Internet of Production: Fundamentals, Methods and Applications. Interdisciplinary Excellence Accelerator Series, Springer (2023). `https://doi.org/10.1007/978-3-031-44497-5{_}3`
30. Morelli, A., Tortonesi, M., Stefanelli, C., Suri, N.: Information-Centric Networking in next-generation communications scenarios. J. Comput. Appl. **80** (2017). `https://doi.org/10.1016/j.jnca.2016.12.026`
31. Özkanlisoy, Ö., Akkartal, E.: The Effect of Suez Canal Blockage on Supply Chains. Dokuz Eylül Üniversitesi Denizcilik Fakültesi Dergisi **14**(1) (2022). `https://doi.org/10.18613/deudfd.933816`
32. Patel, Y.S., Townend, P.: A Stable Matching Approach to Energy Efficient and Sustainable Serverless Scheduling for the Green Cloud Continuum. In: IEEE SOSE (2024). `https://doi.org/10.1109/SOSE62363.2024.00010`
33. Pennekamp, J.: Secure Collaborations for the Industrial Internet of Things. Ph.D. thesis, RWTH Aachen University (2024). `https://doi.org/10.18154/RWTH-2024-03585`
34. Pennekamp, J., Bader, L., Thevaraj, E., Berninger, S., et al.: PRepChain: A Versatile Privacy-Preserving Reputation System for Dynamic Supply Chain Environments. Future Gener. Comput. Syst. **175** (2026). `https://doi.org/10.1016/j.future.2025.108024`
35. Pennekamp, J., Belova, A., Bergs, T., Bodenbenner, M., et al.: Evolving the Digital Industrial Infrastructure for Production: Steps Taken and the Road Ahead. In: Internet of Production: Fundamentals, Methods and Applications. Interdisciplinary Excellence Accelerator Series, Springer (2023). `https://doi.org/10.1007/978-3-031-44497-5{_}2`
36. Pennekamp, J., Buchholz, E., Dahlmanns, M., Kunze, I., et al.: Collaboration is not Evil: A Systematic Look at Security Research for Industrial Use. In: LASER (2021). `https://doi.org/10.14722/laser-acsac.2020.23088`
37. Pennekamp, J., Buchholz, E., Lockner, Y., Dahlmanns, M., et al.: Privacy-Preserving Production Process Parameter Exchange. In: ACSAC (2020). `https://doi.org/10.1145/3427228.3427248`
38. Pennekamp, J., Glebke, R., Henze, M., Meisen, T., et al.: Towards an Infrastructure Enabling the Internet of Production. In: IEEE ICPS (2019). `https://doi.org/10.1109/ICPHYS.2019.8780276`
39. Pennekamp, J., Henze, M., Schmidt, S., Niemietz, P., et al.: Dataflow Challenges in an *Internet* of Production: A Security & Privacy Perspective. In: ACM CPS-SPC (2019). `https://doi.org/10.1145/3338499.3357357`
40. Pennekamp, J., Matzutt, R., Klinkmüller, C., Bader, L., et al.: An Interdisciplinary Survey on Information Flows in Supply Chains. ACM Comput. Surv. **56**(2) (2024). `https://doi.org/10.1145/3606693`
41. Pennekamp, J., Weiler, P., Bodenbenner, M., Sudmann, M., et al.: ConfMod: A Simple Modeling of Confidentiality Requirements for Inter-Organizational Data Sharing. In: MFI5.0 (2025). `https://doi.org/10.1109/NOMS57970.2025.11073722`
42. Poepsel-Lemaitre, R., Kiefer, M., von Hein, J., Quiané-Ruiz, J.A., Markl, V.: In the Land of Data Streams where Synopses are Missing, One Framework to Bring Them All. Proc. VLDB Endow. **14**(10) (2021). `https://doi.org/10.14778/3467861.3467871`
43. Pourbafrani, M., van Zelst, S.J., van der Aalst, W.M.P.: Supporting Automatic System Dynamics Model Generation for Simulation in the Context of Process Mining. In: BIS (2020). `https://doi.org/10.1007/978-3-030-53337-3_19`

44. Rafiei, M., van der Aalst, W.M.P.: An Abstraction-Based Approach for Privacy-Aware Federated Process Mining. IEEE Access **11** (2023). `https://doi.org/10.1109/ACCESS.2023.3263673`
45. Schuh, G., Mauß, W., Potente, T., Schmitz, S., et al.: Green Re-Assembly Upgrade Factory. In: Proceedings of the 31st Aachener Machine Tool Colloquium (AWK '23) (2023). `https://doi.org/10.24406/publica-952`
46. Stoica, I., Morris, R., Karger, D., Kaashoek, M.F., Balakrishnan, H.: Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. ACM SIGCOMM Comput. Commun. Rev. **31**(4) (2001). `https://doi.org/10.1145/964723.383071`
47. Stolz, T., Koren, I., Tirpitz, L., Geisler, S.: GALOIS: A Hybrid and Platform-Agnostic Stream Processing Architecture. In: BiDEDE (2023). `https://doi.org/10.1145/3579142.3594287`
48. Sun, D., Hu, J., Wu, H., Wu, J., et al.: A Comprehensive Survey on Collaborative Data-access Enablers in the IIoT. ACM Comput. Surv. **56**(2) (2023). `https://doi.org/10.1145/3612918`
49. Tacke Genannt Unterberg, L., Koren, I., van der Aalst, W.M.P.: Maximizing Reuse and Interoperability in Industry 4.0 with a Minimal Data Exchange Format for Machine Data. In: Modellierung (2024). `https://doi.org/10.18420/modellierung2024_011`
50. Tirpitz, L.: Towards FAIR Data Stream Processing Ecosystems. In: ACM DEBS (2024). `https://doi.org/10.1145/3629104.3672434`
51. Tirpitz, L., Gentges, L.: Process Model-based Access Control Policies for Cross-Organizational Data Sharing. In: QBD (2024)
52. Tirpitz, L., Kunze, I., Niemietz, P., Gerhardus, K., et al.: Reducio: Data Aggregation and Stability Detection for Industrial Processes Using In-Network Computing. In: ACM DEBS (2025). `https://doi.org/10.1145/3701717.3730547`
53. van der Aalst, W.M.P.: Process Cubes: Slicing, Dicing, Rolling Up and Drilling Down Event Data for Process Mining. In: AP-BPM (2013). `https://doi.org/10.1007/978-3-319-02922-1_1`
54. van der Aalst, W.M.P.: Process Mining: Data Science in Action. Springer, 2nd edn. (2016). `https://doi.org/10.1007/978-3-662-49851-4`
55. van der Aalst, W.M.P.: Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data. In: SEFM (2019). `https://doi.org/10.1007/978-3-030-30446-1_1`
56. van der Aalst, W.M.P.: Federated Process Mining: Exploiting Event Data Across Organizational Boundaries. In: IEEE SMDS (2021). `https://doi.org/10.1109/SMDS53860.2021.00011`
57. van der Aalst, W.M.P.: Object-Centric Process Mining: Unraveling the Fabric of Real Processes. Mathematics **11**(12) (2023). `https://doi.org/10.3390/math11122691`
58. van der Aalst, W.M.P., Jarke, M., Koren, I., Quix, C.: Digital Shadows: Infrastructuring the Internet of Production. In: Internet of Production: Fundamentals, Methods and Applications. Interdisciplinary Excellence Accelerator Series, Springer (2023). `https://doi.org/10.1007/978-3-031-44497-5{_}25`
59. Wiesner, P., Behnke, I., Scheinert, D., Gontarska, K., Thamsen, L.: Let's Wait Awhile: How Temporal Workload Shifting Can Reduce Carbon Emissions in the Cloud. In: Middleware (2021). `https://doi.org/10.1145/3464298.3493399`
60. Zhang, Y., Deng, R.H., Xu, S., Sun, J., et al.: Attribute-based Encryption for Cloud Computing Access Control: A Survey. ACM Comput. Surv. **53**(4) (2021). `https://doi.org/10.1145/3398036`