Received: 23 April 2022 | Revised: 15 October 2024 | Accepted: 17 October 2024

DOI: 10.1111/isj.12567

SPECIAL ISSUE ARTICLE

WILEY

Orchestration logics for artificial intelligence platforms: From raw data to industry-specific applications

Michael Weber [1] | Andreas Hein [2] | Jörg Weking [3] | Helmut Krcmar [1]

[1]Technical University of Munich School of Computation, Information and Technology, Munich, Germany

[2]University of St. Gallen, Institute of Information Systems and Digital Business, St. Gallen, Switzerland

[3]Queensland University of Technology School of Management, Brisbane, Australia

Correspondence
Michael Weber, Technical University of Munich School of Computation, Information and Technology Munich, Munich, Germany.
Email: mic.weber@tum.de

Abstract

Artificial intelligence (AI) platforms face distinct orchestration challenges in industry-specific settings, such as the need for specialised resources, data-sharing concerns, heterogeneous users and context-sensitive applications. This study investigates how these platforms can effectively orchestrate autonomous actors in developing and consuming AI applications despite these challenges. Through an analysis of five AI platforms for medical imaging, we identify four orchestration logics: platform resourcing, data-centric collaboration, distributed refinement and application brokering. These logics illustrate how platform owners can verticalize the AI development process by orchestrating actors who co-create, share and refine data and AI models, ultimately producing industry-specific applications capable of generalisation. Our findings extend research on platform orchestration logics and change our perspective from boundary resources to a process of boundary processing. These insights provide a theoretical foundation and practical strategies to build effective industry-specific AI platforms.

KEYWORDS

AI platform, artificial intelligence, platform orchestration, platform ecosystem

1 | INTRODUCTION

Artificial intelligence (AI) platforms provide organisations with tools and services to use existing or create new AI applications (Geske et al., 2021). These platforms are particularly relevant in industry-specific settings, where firms face structural innovation and market failures (Haki et al., 2024; Jacobides, Cennamo, & Gawer, 2024). AI platforms address these problems by orchestrating autonomous actors in developing and consuming applications (Geske et al., 2021; Jacobides et al., 2021). For instance, in manufacturing, AI platforms offer standardised interfaces for integration with heterogeneous equipment providers (Pauli et al., 2021). In healthcare, they facilitate collaboration among hospitals and experts to improve AI models for diagnosis. This collaborative approach leads to more accurate and robust models for diagnosis—outcomes that organisations might struggle to achieve independently (Thrall et al., 2018).

The success of AI platforms in industry-specific settings hinges on effective platform orchestration, defined as the coordination of autonomous actors' efforts and resource contributions (Tiwana, 2014) to navigate the complex challenges unique to these environments. For instance, developing an AI application to diagnose specific diseases from medical scans requires unique datasets and domain-specific expertise (Rajpurkar et al., 2022). These resources are typically scarce and dispersed among actors, with data privacy and intellectual property complicating resource exchanges (Cohen & Mello, 2019). Additionally, industry-specific AI applications struggle to generalise to heterogeneous users without costly retraining (Brecker et al., 2023). For example, variations in production setups can negatively impact the performance of applications trained in different manufacturing contexts (Weber et al., 2022). This limited generalizability discourages producers and risks suboptimal performance of context-sensitive AI applications (Brecker et al., 2023).

The literature on platform orchestration primarily focused on homogenous markets with standardised products and services (De Reuver et al., 2018; Rietveld & Schilling, 2021). This literature highlights two dominant orchestration logics (Cusumano et al., 2019; Evans & Gawer, 2016): promoting third-party innovation through boundary resources to extend the platform (Eaton et al., 2015; Ghazawneh & Henfridsson, 2013) and facilitating transactions in multi-sided markets (Rochet & Tirole, 2003; Shi, 2023). While helpful in explaining orchestration in homogeneous markets, this focus leaves a research gap regarding the challenges of industry-specific AI platforms, which must navigate environments characterised by specialised resources, data-sharing concerns, high user heterogeneity and context-sensitive applications. These challenges necessitate new orchestration strategies to realise the potential of industry-specific AI platforms.

In light of these challenges, AI platforms like Hugging Face and Nvidia have developed novel orchestration logics that could help better understand how industry-specific AI platforms can mitigate these problems. Hugging Face employs a community-driven approach, orchestrating actors to build and share pretrained models for further refinement, thus catering to a broad range of third-party developers' needs. In addition, Nvidia facilitates federated learning in specific industries, enabling multiple actors to build more generalizable models that can serve heterogeneous users (Roth et al., 2022). Despite these practical advancements, we do not understand how these approaches mitigate the orchestration challenges, highlighting a significant gap in the current literature. Consequently, we investigate how different platform orchestration logics can address orchestration challenges in industry-specific AI development.

We employed a multiple-case study approach (Eisenhardt, 1989; Yin, 2018) in the medical imaging sector to explore the orchestration logics of industry-specific AI platforms. A multiple case study allows us to compare different orchestration logics and their impact on addressing the orchestration challenges in industry-specific AI development. We selected this industry due to its maturity of AI platforms and the pronounced orchestration challenges, such as the scarcity of specialised data and domain knowledge, strict data privacy regulations, and high user heterogeneity. The study encompassed five cases based on 38 interviews and archival data, analysed using inductive coding methods (Gioia et al., 2013; Strauss & Corbin, 1990). Our selection criteria included AI platforms in medical imaging with different orchestration logics, mature products, and an active user base.

The results reveal four orchestration logics that AI platforms apply and combine: platform resourcing, data-centric collaboration, distributed refinement, and application brokering. The logic of platform resourcing, exemplified by the platform DataForge, provides foundational resources, such as developer services, to enable a broad spectrum of actors to mobilise their data and expertise for AI development. The logic of data-centric collaboration, exemplified by RadiaHub, facilitates multiple actors' co-creation of datasets and models, yielding more diverse datasets and generalizable models. The logic of distributed refinement, exemplified by ModelCraft, describes actors' distributed sharing and refinement of datasets and models via the platform, yielding increasingly specialised resources (e.g., fine-tuned models). Last, the logic of application brokering, exemplified by MedConnect, facilitates matchmaking through consumers' benchmarking of context-specific applications.

Our findings have implications for research on platform orchestration in general and AI platforms in particular. While extant literature has focused on traditional innovation and transaction logics (Cusumano et al., 2019; Evans & Gawer, 2016), our study suggests that platform orchestration can be more complex, requiring additional logics to cope with the distinct challenges in industry-specific AI development. We also extend the concept of boundary resources on AI platforms (Ghazawneh & Henfridsson, 2013), showing how autonomous actors process and refine data and models as ecosystem resources at the boundary. This boundary processing requires decoupling data and model resources from the platform, which platform owners should consider in their architecture design. Last, we provide practical implications, the limitations of this study, and avenues for future research.

2 | BACKGROUND AND RELATED WORK

This section introduces AI and outlines the necessary resources and challenges organisations face when developing AI applications. Next, we present AI platforms and synthesise research on platform orchestration. By integrating these insights, we highlight the orchestration challenges unique to industry-specific AI platforms, demonstrating how our current understanding of platform orchestration falls short in addressing these complexities.

2.1 | AI development

AI represents a broad and continuously evolving research field that aims to make machines intelligent (Russell & Norvig, 2021). AI comprises numerous techniques, such as machine learning (ML), reasoning, computer vision, and natural language processing (Russell & Norvig, 2021; Stone et al., 2016). The current wave of AI centers largely around applications that use ML techniques, particularly deep learning (Berente et al., 2021; Stone et al., 2016). ML techniques allow systems to learn from experience and automatically improve their performance on a given task (Jordan & Mitchell, 2015). In line with this outlook, we refer to AI development as the development of applications based on ML.

For AI development, organisations typically engage in an iterative process of data management, model training, model verification and model deployment (Figure 1). Throughout this process, organisations create a variety of artefacts, from training datasets to deployable applications (Table 1).

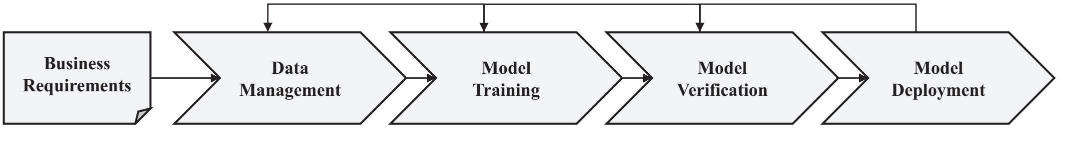

FIGURE 1 Artificial intelligence development process (adapted from Ashmore et al. (2021)).

TABLE 1 Artefacts of AI development (adapted from Duda et al. (2023)).

Artefact	Definition
Training dataset	A dataset that is used to train the model. In supervised learning tasks, training data typically needs to be annotated by humans (i.e., assigned labels which are to be inferred).
Verification dataset	A dataset that is used to verify trained models before deployment. The verification dataset should mirror the data distribution of its intended usage context or allow testing for edge cases and fairness.
Input data	Data that is typically gathered in real-world processes and sent to the model to perform inferences after deployment.
Trained model	A model derived from ML that can infer an outcome based on input data (e.g., predict the weather for the next day).
Pretrained model	A trained model that is used as a foundation for model training in other contexts. For example, a pretrained model can be fine-tuned on new training data to adapt to a different task.
AI application	Comprises a deployable model alongside complementary software to integrate the model into existing systems and handle model requests.

Abbreviations: AI, artificial intelligence; ML, machine learning.

AI development requires substantial resources, data and expertise, which most firms do not possess. First, organisations require data in sufficient quantity and quality for model training and verification. Most firms find this data difficult to access, biased or unbalanced, or costly to prepare (Jöhnk et al., 2021; Pumplun et al., 2019). Second, organisations require expertise for data labeling and model training, which is typically scarce and costly (Jöhnk et al., 2021; van den Broek et al., 2021). Third, training models require vast computational resources and access to a dedicated infrastructure (Jöhnk et al., 2021; Lins et al., 2021). These resource requirements become especially relevant for larger models that use deep learning techniques and require multiple rounds of training. Fourth, applications' actual deployment and operation come with significant additional costs of integration into real processes and legacy systems (Paleyes et al., 2022). These costs further increase due to required investments into the ongoing monitoring, evaluation, and retraining of models, as their performance might change over time (Paleyes et al., 2022; Weber et al., 2023).

Thus, organisations must access and integrate diverse resources to manage the intricacies of AI development. Many organisations lack these resources and thus rely on collaboration with external actors to overcome these challenges.

2.2 | AI platforms and platform orchestration

Following Geske et al. (2021, p. 12), AI platforms present a specific type of platform that "*provide[s] organisations with access to AI technology to support them in creating or using AI applications through federating and coordinating constitutive agents, leveraging value by enabling economies of scope, and entailing modular technological architecture*." As such, AI platforms orchestrate autonomous actors by supporting them in developing custom AI applications and providing access to readily usable applications (Geske et al., 2021; Lins et al., 2021).

Given the nascent development of research on orchestration in AI platforms (H. Li & Kettinger, 2021; Rai et al., 2019), we first draw on general platform orchestration research to illustrate innovation and transaction as the two dominant orchestration logics (Cusumano et al., 2019; Evans & Gawer, 2016). These two logics present structural solutions on how platforms guide autonomous actors' efforts and resource contributions to address innovation and market failures (Jacobides et al., 2024). Based on these insights, we synthesise literature on AI platforms (Geske et al., 2021).

Innovation platforms such as Apple's iOS exemplify the orchestration logic of innovation where the platform owner provides resources to stimulate third-party developments that extend the platform (Tiwana, 2014). These platforms foster innovation by converging technological standards, facilitating knowledge exchanges, and incentivising participation in the ecosystem (Jacobides et al., 2018; Jacobides et al., 2024). To support these activities, platform owners develop and offer boundary resources, including software development kits (SDKs) and application programming interfaces (APIs) (Engert et al., 2023; Foerderer et al., 2019; Hein et al., 2019). The platform owner provides and curates these boundary resources (Ghazawneh & Henfridsson, 2013), although being influenced by distributive dynamics in the ecosystem (Eaton et al., 2015).

Innovation platforms ensure that the applications and modules developed by third parties are compatible with the platform through standardisation and decoupling (Tiwana, 2014; Tiwana et al., 2010). By providing standardised boundary resources, innovation platforms can ensure the interoperability of third-party applications and modules with the platform (Baldwin & Woodard, 2009). Furthermore, decoupling modules from the platform core fosters innovation through the combination and recombination of modules while ensuring the stability of the platform core (Benlian et al., 2018; Tiwana et al., 2010).

Transaction platforms such as Airbnb exemplify the orchestration logic of transaction where the platform orchestrates transactions between two or more market sides. These platforms are intermediaries, enabling efficient exchanges between autonomous actors, reducing transaction costs, and enhancing market efficiency (Rochet & Tirole, 2003; Shi, 2023). The key to their operation is managing and optimising the volume and quality of transactions to ensure a seamless and value-adding experience for all participants (Eisenmann et al., 2006). To that end, platform owners rely on algorithms to provide efficient search and matchmaking. Additionally, platform owners employ governance mechanisms to assess the credibility and quality of products or services (Boudreau, 2010; Wareham et al., 2014).

Literature on AI platforms primarily focused on delineating the different types of AI platforms and the boundary resources they provide (Geske et al., 2021; Lins et al., 2021). While some AI platforms are industry-agnostic, other AI platforms cater to industry-specific needs to support tailored applications within an industry (Geske et al., 2021). The literature further highlights how AI platforms range from supporting AI development to offering ready-to-use applications (Geske et al., 2021; Jacobides et al., 2021). The support in AI development follows the innovation logic from general platform research. Here, the platform owner provides software services (e.g., automated ML), developer services (e.g., SDKs, data labeling tools), and infrastructure services (e.g., computing resources, data storage) to enable autonomous actors to develop their own applications (Jacobides et al., 2021; Lins et al., 2021). In addition, some AI platforms provide access to readily usable AI applications in homogenous markets (Brecker et al., 2023; Diaferia et al., 2022), following the traditional transaction orchestration logic.

2.3 | Orchestration challenges for industry-specific AI platforms

However, beyond these initial insights into the general innovation and transaction logics, our understanding of how AI platforms can orchestrate autonomous actors in more complex, industry-specific settings remains limited: These settings face challenges related to specialised resources, data-sharing concerns, high heterogeneity of users, and context-sensitive applications. Prior research on platforms with similar constraints, such as in the Industrial Internet of Things (IIoT), demonstrates that without effective orchestration, these challenges can lead to endogenous platform failure (Pauli et al., 2021; Wlcek et al., 2023).[1] We detail these orchestration challenges of industry-specific AI platforms below.

[1]Examples in the context of IIoT, such as GE's Predix and Siemens Mindsphere, illustrate that these industry-specific challenges still present unresolved orchestration challenges for firms.

The first industry-specific challenge is providing access to specialised resources to drive innovation (Haki et al., 2024). For example, AI development in manufacturing requires highly specialised datasets for a broad spectrum of use cases, such as predictive maintenance and machine optimization. Unlike traditional innovation platforms, the platform owner cannot supply all the required datasets, labeling services or pretrained models (Haki et al., 2024). This limitation restricts innovation to a few powerful actors who can provide these resources on their own (Jacobides et al., 2021). As a possible solution, we observe how AI platforms such as HuggingFace employ a community-driven approach, where third parties contribute pretrained models to the platform, which, in turn, could cater to a broad range of specialised use cases.

The second industry-specific challenge is addressing data-sharing concerns, which hinder collaboration and innovation on AI platforms. This issue is particularly prominent in industries where organisations are reluctant to share data due to competition, security, privacy concerns, and legal compliance (Jussen et al., 2024). For example, in healthcare, data is often subject to privacy concerns and strict regulations such as HIPAA (Cohen & Mello, 2019). The reluctance to share data limits innovation and can result in biased or unbalanced datasets, reducing the platform's overall value. While traditional innovation platforms do not account for these data-sharing challenges, research on data platforms and ecosystems suggests mechanisms such as data security, encryption, decentralised storage, and trust-building (Jussen et al., 2024; Otto & Jarke, 2019). However, if and how novel orchestration approaches, such as Nvidia's federated learning approach (Roth et al., 2022), can potentially address data-sharing concerns on industry-specific AI platforms remains to be understood.

The third industry-specific challenge is addressing the heterogeneity of user requirements on AI platforms. In manufacturing, for instance, production setups and systems can vary significantly across users (Pauli et al., 2021). This heterogeneity could disincentivize producers to extend the platform, as integration becomes costly and the number of suitable consumers for standardised applications is limited. For example, an AI application optimised for one specific machine might not generalise to users with different machines, production setups, or data distributions (Weber et al., 2022). In addition, customising AI applications for varied contexts is often complex or even unfeasible, as it requires retraining or significant adjustments to the model (Brecker et al., 2023; Diaferia et al., 2022). Traditional innovation platforms typically cater to more homogeneous user groups with standardised products or services, making them less equipped to handle the heterogeneity found in industry-specific settings (Pauli et al., 2021).

The fourth industry-specific challenge is the context sensitivity of AI applications (Brecker et al., 2023), which complicates their universal quality assessment and hinders effective matchmaking in multi-sided markets. Determining an AI application's quality is not trivial and highly dependent on its fit with a consumer's unique data environment (Paleyes et al., 2022). As producers typically aim to develop standardised applications, consumers risk unknowingly accepting suboptimal performance given their unique data environments (Brecker et al., 2023; Diaferia et al., 2022). This suboptimal performance, in turn, not only affects an organisation's operations but can also lead to ethical and compliance issues (Lins et al., 2021). The transaction logic from general platform research proposes various mechanisms, such as recommender algorithms, user ratings, and certification programs, to signal the quality of applications and facilitate transactions (Haki et al., 2024). However, these generic mechanisms assume that the quality of an application can be universally determined, not accounting for an AI application's performance in specific consumer contexts.

In conclusion, existing literature highlights that industry-specific AI platforms face orchestration challenges that traditional platform logics cannot fully address. However, anecdotal evidence from platforms like Hugging Face and Nvidia suggests that new orchestration strategies may offer solutions to these challenges.

3 | METHOD

To explore the specific aspects of platform orchestration in the context of industry-specific AI platforms, we follow a multiple-case study focusing on five distinct AI platforms in medical imaging (Eisenhardt, 1989; Yin, 2018). This method is appropriate as it allows for an in-depth exploration of different orchestration logics that require context-sensitive knowledge of a relatively unexplored phenomenon (Siggelkow, 2007). Furthermore, a cross-case analysis allows us to examine orchestration logics across different platforms, thereby uncovering patterns and variations (Eisenhardt & Graebner, 2007). Last, we have chosen the medical imaging sector due to its growing AI market (van Leeuwen et al., 2021), diverse range of mature platforms for analysis, and rich context to examine the intricacies of platform orchestration (Hosny et al., 2018; Thrall et al., 2018).

We collected and analysed 38 interviews and archival data, employing inductive coding methods (Gioia et al., 2013; Miles et al., 2018). This rich data set, encompassing various actor roles, such as platform owners, complementors and customers, was crucial in understanding how different AI platforms orchestrate autonomous actors' efforts and resource contributions. Our results offer valuable insights into their orchestration logics based on an inductive coding of each case and a subsequent cross-case analysis (Eisenhardt & Graebner, 2007; Strauss & Corbin, 1990).

3.1 | Data collection

Our case sampling followed four selection criteria to answer the research question. First, we sampled AI platforms with different motives and orchestration logics, initially guided by our pre-understanding of the two dominant orchestration logics of innovation and transaction (Jacobides et al., 2024). Second, we only considered AI platforms that facilitate AI development for medical imaging to ensure comparability between cases. Third, we only considered AI platforms with mature products and an active user base to ensure the robustness of our results. Fourth, we only considered AI platforms operating in Northern America and Europe to control for legal and cultural factors.

Based on these criteria, we conducted an initial round of data collection from 2019 to early 2020, conducting 17 interviews subject to the DataForge, RadiaHub and ClinDeploy AI platforms. We also collected archival data such as website information, documentation, and whitepapers to ensure data triangulation (Yin, 2018). The first round aimed to explore AI platforms in the field, focusing on the platform's offerings, the actors involved, and the key challenges to AI development. During this round, we found that AI platforms employ multiple orchestration logics, and some aspects appeared to differ from non-AI platforms. For example, we found that AI platforms seemed to apply a novel type of orchestration logic that involves peer-to-peer collaborations and aims to address challenges related to data diversity.

Informed by the results of the first round, we engaged in a second round of data collection from 2021 to 2023, where we collected additional data and added the cases ModelCraft and MedConnect. We did this to learn more about the emerging orchestration logics and their motives, variations, and inherent challenges. During this stage, our focus shifted to the AI-specific aspects of platform orchestration (e.g., the role of datasets, federated learning, and the context sensitivity of AI applications). We selected the cases ModelCraft and MedConnect for theorization, as they allowed us to study the emerging orchestration logics in different settings to capture contextual nuances and increase the robustness of our findings. During this round, we collected 21 interviews and additional archival data. After analysing the data, we realised the last five interviews did not provide novel insights relevant to our research question. Hence, we were confident that theoretical saturation was reached, and we concluded our data collection (Strauss & Corbin, 1990).

To conduct the interviews in both rounds, we used semi-structured interview guidelines and followed the suggestions provided by Myers and Newman (2007). We conducted 33 interviews online using video conferencing tools and five in person. We recorded and transcribed all interviews for the coding process. We sampled our interview partners to cover different roles and perspectives of the platform owner as the central actor (e.g., management, engineers and clinical experts). Furthermore, we aimed to cover the roles of platform users (e.g., hospitals, research institutes, vendors) to gain insights into how they interact with the platform and to address potential bias from the

TABLE 2 Overview of the research process.

	Round 1: In-field exploration of AI platforms and orchestration logics	Round 2: Iterative data collection, analysis, and theorising on orchestration logics
Timeline	2019–2020	2021–2023
Focus and key activities	Focus: • Gain an overview of AI platforms • Understand how AI platforms facilitate AI development Key activities: • Basic research on context and platforms as the unit of analysis • Conduct and analyse the first set of interviews • Within and cross-case analysis	Focus: • Gain additional information on emergent orchestration logics • Theorise on orchestration challenges as sources of innovation and market failures in the context of industry-specific AI • Theorise how orchestration logics address orchestration challenges • Reach theoretical saturation Key activities: • Conduct and analyse the second set of interviews • Within and cross-case analysis • Constant comparison with literature
Sampling Strategy	Sampling strategy: • AI platforms with distinct orchestration logics • AI platforms in medical imaging • AI platforms with established market • AI platforms in North America and Europe Interviews in Round 1: • DataForge: 5 Interviews • RadiaHub: 5 Interviews • ClinDeploy: 7 Interviews	Sampling strategy: • Two additional AI platforms in medical imaging to increase robustness and capture contextual nuances Interviews in Round 2: • DataForge: 2 Interviews • RadiaHub: 1 Interview • ModelCraft: 7 Interviews • ClinDeploy: 2 Interviews • MedConnect: 9 Interviews
Insights	• AI platforms employ multiple orchestration logics to address challenges in AI development • Peer-to-peer collaboration and the sharing of data and models present conceptually distinct types	• AI platforms apply four distinct orchestration logics to address orchestration challenges in industry- specific AI development • Key orchestration challenges include the need for specialised resources and the heterogeneity of user requirements

Abbreviation: AI, artificial intelligence.

platform owner's perspective. Table A1 provides an overview of all interviews, and Table A2 and Table A3 show the principal interview guidelines used for data collection.

Table 2 summarises the two rounds of data collection, the respective timeline, key activities, our sampling strategy, and derived insights.

3.2 | Data analysis

Our data analysis comprised within- and cross-case analysis (Eisenhardt, 1989; Yin, 2018) and followed an iterative coding procedure with increasing levels of abstraction (Gioia et al., 2013; Miles et al., 2018). We started data analysis by familiarising ourselves with the cases and their context by repeatedly reviewing the interviews and archival data sources. We then constructed case descriptions for each platform to understand its value proposition, offerings, relevant actors and their motives for participation. We then discussed and refined these descriptions within the research team until we understood each platform robustly. Table A4 gives a brief description of the AI platforms included in this case study.

We then systematically coded the interviews and archival data (Miles et al., 2018; Strauss & Corbin, 1990). Following our research question on platform orchestration, we initially sought to understand how the platform guides different actors and how they engage with and through the platform. Therefore, two researchers started coding instances of actors engaging with the platform. Guided by our conceptual understanding of platform orchestration, we especially looked for instances of actors using platform resources for AI development, resource contributions to the platform (e.g., third-party applications), and transactions mediated by the platforms. Along with that, we coded the involved actor roles (including the platform owner), their drivers and motives (e.g., data access), the perceived outcomes (e.g., more diverse datasets), and the perceived challenges of those platform engagements.

We then grouped the lower-level codes into categories (Gioia et al., 2013) to derive more general patterns of platform engagements (e.g., contributing datasets to the platform). In line with our research goals, we decided to focus our analysis on patterns of platform engagement that are immediately concerned with AI development (e.g., datasets, model training). For example, we did not further consider more generic instances, such as requesting platform features or providing feedback to the platform owner, as we would not expect to find novel facets in the context of industry-specific AI development here. We sorted and clustered the categories to gain a robust understanding of each pattern, especially its mechanics, actors, motives, outcomes and challenges. At this stage, we started comparing the categories across cases to refine the categories further and arrive at consistent meanings (Miles et al., 2018). We further used the AI development process model (Ashmore et al., 2021) as a tool to understand better how each emergent pattern relates to the process of AI development.

We then engaged in a theorising process to understand the overarching orchestration logics employed by AI platforms and answer our research question. For this, we build on our theoretical understanding of platform orchestration as a theoretical lens, which suggests that platform orchestration coordinates autonomous actors' efforts and resources to overcome orchestration challenges that can otherwise lead to innovation or market failures (Jacobides et al., 2024). Hence, we looked at the different platform engagement patterns and analysed the underlying orchestration challenges they aim to solve in the context of industry-specific AI development. During this stage, we again compared the emerging relations across cases to confirm, extend and sharpen our conclusions (Eisenhardt, 1989) and constantly compared our findings with general platform research.

After several rounds of going back and forth, this process led us to derive four theoretically distinct orchestration logics that AI platforms apply (in isolation or combination): platform resourcing, data-centric collaboration, distributed refinement, and application brokering. We then described the orchestration logics and relevant variations. Our descriptions focus on the solutions provided by the platform, the actors involved, and the respective outcomes. Last, we contrasted the empirically derived orchestration logics with conventional insights on platform orchestration (e.g., the traditional innovation logic) to derive theoretical implications.

Figure 2 provides an overview of the coding procedure.

4 | EMPIRICAL FINDINGS

We identified four distinct orchestration logics for AI platforms in medical imaging: platform resourcing, data-centric collaboration, distributed refinement and application brokering. This section details the orchestration logics and the challenges they address. Table 3 provides a more comprehensive cross-case comparison, illustrating how the cases apply the orchestration logics.

4.1 | Platform resourcing

As part of the orchestration logic of platform resourcing, AI platforms provide a modular technological core and supporting resources that enable individual actors to create new datasets, models, and applications that extend the

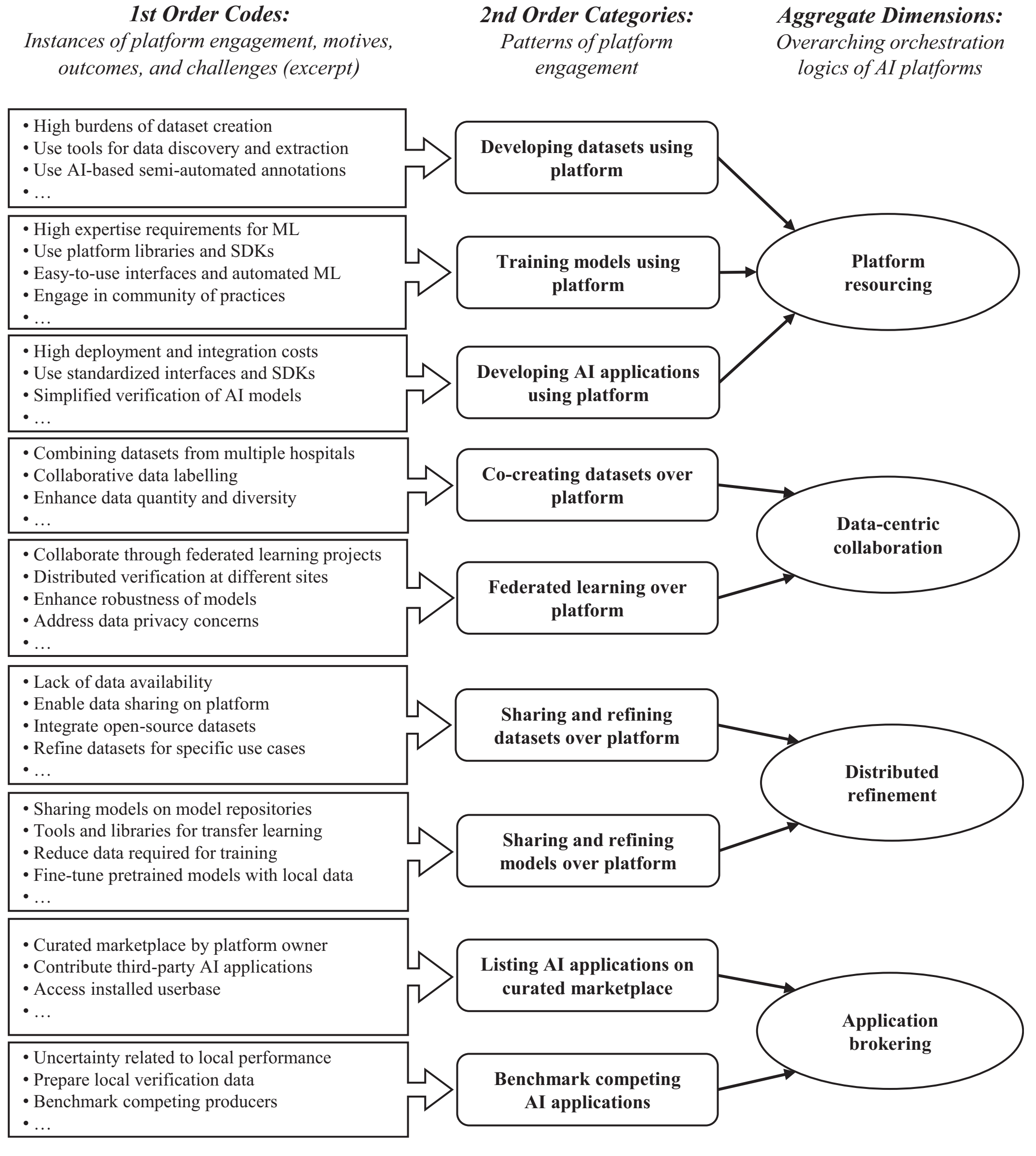

FIGURE 2 Stylized illustration of coding procedure (according to Gioia et al. (2013)). AI, artificial intelligence; ML, machine learning; SDKs, software development kits.

platform. The supporting resources provided by the platform owner comprise a range of software, developer, and infrastructure services for AI development. Whereas platforms like DataForge, RadiaHub, and ModelCraft focus on providing resources for the data management and model training phases, platforms like ClinDeploy and MedConnect focus on model verification and deployment. A broad range of actors engage through this logic, including hospitals, radiology centers, individual developers, researchers, and commercial vendors. Consequently, the motives range from experimenting with AI to conducting scientific research and developing internal and commercial applications.

AI platforms aim to lower the cost and expertise required for AI development through this orchestration logic. Platforms aim to achieve this by providing domain-specific tools that automate resource-intensive steps and abstract

TABLE 3 Cross-case comparison of orchestration logics.

	Platform DataForge	Platform RadiaHub	Platform ModelCraft	Platform ClinDeploy	Platform MedConnect
Platform resourcing	• Easy-to-use tools for data extraction, labeling, and semi-automated annotations • Use a medical viewer for model verification • Integrate with Python libraries for model training	• Easy-to-use tools for data extraction and labeling • Use automated ML features • Build on standardised use case descriptions and interfaces • Code-level access for customizability	• Use domain-specific libraries and software development kits • Use tools for labeling and semi-automated annotation • Connect to scalable computing infrastructure • Code-level access for customizability	• Use tools for data extraction • Use well-defined interfaces for deployment • Integrate easily through an operating system • Tools to verify and benchmark the performance of models	• Use well-defined interfaces for deployment • Integrate easily through an operating system • Tools to verify and benchmark the performance of models • Access scalable cloud infrastructure
Data-driven collaboration	• Combine datasets in public or private data projects • Collaboratively annotate data • Integrate with third-party federated learning frameworks	• Participate in collaborative data projects on the platform • Use federated learning framework • Use connected users for distributed verification	• Use federated learning framework • Use connected users for distributed verification	• Use connected users for distributed verification and continuous monitoring	• Use connected users for continuous monitoring
Distributed refinement	• Share datasets on the platform publicly or privately • Compare and use third-party datasets for model training • Share pretrained models on the platform publicly or privately • Compare and use third-party pretrained models	• Share datasets on the platform publicly or privately • Compare and use third-party datasets for model training • Share pretrained models on the platform publicly or privately • Compare and use third-party pretrained models	• Share pretrained models in the platform repository • Access third-party pretrained models for AI development • Use libraries for transfer learning		
Application brokering				• Contribute AI applications to a curated marketplace • Access and compare third-party AI applications	• Contribute AI applications to a curated marketplace • Access and compare third-party AI applications

Abbreviation: ML, machine learning.

complexities. Furthermore, platforms offer seamless integration and connectivity with medical information systems to reduce deployment costs. Thereby, AI platforms also enable more resource-constrained actors to mobilise their data and expertise and develop applications. Besides, platforms emphasise interoperability with external developer tools to provide full flexibility for AI development.

For example, platforms like DataForge provide data management services that facilitate data collection, retrieval, analysis, and preparation needed in model training and verification. DataForge provides tools to sample, import, and anonymize medical images from medical information systems. Once imported, DataForge provides tools that allow domain experts to label medical data, emphasising usability and convenience. Furthermore, interfaces to other developer tools (e.g., Python libraries) provide full flexibility for subsequent model training and deployment. Stakeholders from DataForge's platform report how the platform addresses some important barriers in AI development, specifically those tied to data management:

> "*It is common sense in data science that 80% of the time you spend curating data, like cleaning and annotating, and just 20% of the time, you are training models, right? So, if most of the effort and time is spent working on the data before training the models, we must have very good tools to do that*. […] *So, I hadn't yet found an annotation tool to do [multi-level annotation] specifically for medical images that could handle DICOM images. In my opinion, the one with almost everything you need to annotate radiology images is [DataForge]*."
>
> (AI engineer and radiologist, DataForge user)

> *"The annotation process is obviously very critical, but we also have tools where engineers can access the annotations in real-time, and then they can either build an algorithm or, what's becoming more popular is, let's say, you have a huge data set of 100,000 x-rays, and you want to annotate all 100,000 now. Ideally, you don't have to get a human expert to do all that. That's very time-consuming and expensive. So, what some of our users do is they can see the annotations and build an algorithm to help annotate the rest of the dataset. You can bootstrap it and then accelerate the annotation."*
>
> (Medical advisor, DataForge)

As another example, platforms like ClinDeploy focus on resources for model deployment and verification that facilitate the development of AI applications. Specifically, ClinDeploy provides a cloud-based operating system and SDKs that facilitate model deployment, serving, and monitoring. The operating system provides a standardised interface that integrates with existing medical workflows and information systems, significantly reducing deployment complexity and effort. Actors can flexibly develop applications by combining several models simultaneously within more complex workflows. The stakeholders from ClinDeploy's platform report on the perceived benefits and how it enables AI development on the platform:

> "*[We offer] complete integration, and that [integration] is really complicated for someone who wants to develop models. She doesn't really want to worry about the integration; it is not interesting what kind of images, where, when, and how they are sent. She wants a fixed interface with an input and an output*."
>
> (CTO, ClinDeploy)

> *"I believe that [a platform] solution is probably the only sensible one in the end because when I see how difficult it is for us to make all these installations [for model deployment] and then tell some practice, well, they need ten different systems now, […] and then next year comes the eleventh or so, that's not manageable.*"
>
> (Radiologist and researcher, ClinDeploy user)

In conclusion, AI platforms apply the logic of platform resourcing to provide the modular technological core and supporting resources that enable the development of compatible datasets, models and applications. This logic emphasises reducing the cost and effort required for AI development, allowing a broad spectrum of otherwise resource-constrained actors to mobilise their own data and expertise and participate in AI development.

4.2 | Data-centric collaboration

In the orchestration logic of data-centric collaboration, AI platforms facilitate the collaboration between multiple actors regarding the co-creation of datasets and models through merging datasets, federated learning, and distributed verification. In contrast to the logic of platform resourcing, data-centric collaboration involves multiple actors combing their data for AI development, yielding more diverse and balanced datasets and, ultimately, more generalizable models. While platforms like DataForge, RadiaHub, and ModelCraft support all identified forms of collaboration, platforms like ClinDeploy and MedConnect only support applications' distributed verification and monitoring. The participating actors are mostly peer organisations with access to data, such as hospitals and research institutions, and often aim to achieve mutual benefits from their collaboration. Hence, we label these actors as peer collaborators in AI development.

Through this orchestration logic, AI platforms aim to address the lack of data diversity caused by siloed data and strong data-sharing concerns. To that end, AI platforms aim to reduce the friction from such collaborative engagements, allowing more actors to participate and reap the mutual benefits. AI platforms facilitate collaboration by providing the necessary boundary resources (e.g., federated learning frameworks), promoting standardised interfaces and formats (e.g., model type, data labeling conventions), and facilitating coordination (e.g., acting as a mediating entity). Beyond possible knowledge spillovers within these collaborations, combining multiple actors' diverse data creates a more tangible benefit. Actors create more diverse and balanced datasets based on this collective data diversity and achieve higher generalizability in the resulting models. Such a lack of data diversity could otherwise critically reduce the value of the resulting models and applications.

For example, platforms like DataForge support the merging and preparing of datasets between multiple collaborators. DataForge provides the necessary tooling to import, deanonymize, and merge datasets from multiple actors. Furthermore, DataForge's tools allow multiple actors to work on data labeling collaboratively while ensuring controlled and trusted access to the data. Moreover, DataForge promotes standardised formats that facilitate data integration and processing. Stakeholders from DataForge's platform report how the platform addresses some of the barriers when collaborating on datasets and the benefits it enables:

> "*Our platform also supports the upload of multiple datasets. For example, when a few hospitals collaborate, they can easily organize their data and select what they want to work on. But of course, security and safety are always a concern, so we allow project managers to control user access very flexibly*."
>
> (Business manager, DataForge)

> *"If you pick your institutions properly, if they're diverse enough across the country and diverse enough in terms of the scanner hardware they have, and really if you have 5 or 6 or let's say 5 to 8 institutions that have enough intrinsic data diversity, I think, you would have something that would then work at, you know, maybe upwards of 90 percent or 95 percent of places across the country"*
>
> (Professor of radiology, DataForge user)

As another example, platforms like RadiaHub support federated learning and the distributed verification of models involving multiple collaborators. RadiaHub provides the software and standardised interfaces that participating actors need to install locally to train, verify, and exchange models. Furthermore, RadiaHub actively develops and suggests standards for certain AI use cases that facilitate consensus formation (e.g., output formats and labeling conventions). The main benefit of federated learning is that no data needs to be shared, alleviating data-sharing concerns and incentivising participation. On the other hand, federated learning raises new challenges, such as setting up infrastructure at each site or fairly distributing the value from the jointly created outcome. Stakeholders from RadiaHub report on their experiences with federated learning and distributed verification involving multiple actors:

> "*It turns out everybody's […] anxious to give up their data. So, that was one reason why we developed this federated idea that you bring the algorithm to the [hospital] site rather than the site giving you their data. And that way we can still have the diversity from all over the country […] but without having to pull the data into one place*"
>
> (Vice president data science, RadiaHub)

> "*[…] we think the strength of [RadiaHub] is our connectivity for federated learning, and we've already done a number of federated learning projects. We've built several models through federated learning that have been tested locally where the sites that built them were held out. Now we're using that same network to evaluate those at sites where they didn't participate in the experiment*."
>
> (Chief medical information officer, RadiaHub)

> "*You have test data at each site, and you train this algorithm, move it around, and view its performance at each site […]. But the more interesting thing is getting test data from an institution that the model wasn't even trained at. […] I think that's the real test because the entire point of building a model trained at multiple sites is […] to get to the point where you have [a model] that will be generalizable and work at sites it hasn't been to before because that's how you achieve scale*."
>
> (Professor of radiology, RadiaHub user)

In conclusion, AI platforms apply the logic of data-centric collaboration to enable actors to combine their data in a trusted way, creating more diverse and balanced datasets and more generalizable models across heterogeneous users.

4.3 | Distributed refinement

As part of the orchestration logic of distributed refinement, AI platforms allow actors to share and further refine the previously created datasets and models. Actors thus cultivate specialised resources for AI development that go beyond the platform-provided resources (e.g., pretrained models for specific diseases). Platforms like DataForge, RadiaHub and ModelCraft, which focus on the data management and model training phases, apply this logic. The participating actors vary in type and motive depending on the resource transacted. For example, these actors include data owners and domain experts (e.g., hospitals and medical imaging centers) and model developers (e.g., research institutions, hospitals and commercial vendors). As these actors both consume and produce resources, we refer to these as prosumers in AI development.

Through this orchestration logic, AI platforms incentivise and facilitate the sharing of datasets and models and their subsequent reuse in AI development. AI platforms aim to achieve this through proposing and enforcing standardised formats and interfaces (e.g., defining data specifications), creating transparency of the resources available (e.g., providing repositories), and making external resources reusable and refinable (e.g., using transfer learning frameworks). Through the specialised contributions of external prosumers (e.g., curated datasets for specific modalities), AI platforms can support AI development for a broader range of industry-specific use cases. While the provided datasets address barriers related to data availability, pretrained models significantly reduce the data and computing power required to train a model.

For example, platforms like DataForge facilitate the sharing and distributed refining of datasets through multiple prosumers. To that end, DataForge provides an infrastructure to anonymize and share datasets publicly or with selected actors. Other actors can then search for suitable datasets and use DataForge's tools to annotate the data further, enrich the data with their own data sources, or train new models with the data. Given the sensitive nature of patient data, the shared data often represents open data. While not always perfectly fitting, those open datasets can serve as a valuable starting point for further refinement:

> "*We're using our public datasets. Most of it is from the [public data archive], which is just an amazing resource, and we have quite a lot of it hosted on our site. [...] Mostly, we're using it just to give people a way to access that data and look through it, but we're also annotating on it*."
>
> (Project manager, DataForge)

> *"I would say that we have used some open data from some competitions to build models. Surprisingly, these models generalized very well in our own data without the need for transfer learning or fine-tuning [...]. We also have a model to segment and give the volume of [organ], and for this one, the first training was done with public images, but then we annotated our own images to improve the accuracy."*
>
> (AI engineer and radiologist, DataForge user)

As another example, platforms like ModelCraft facilitate the sharing and distributed refinement of pretrained models. To that end, ModelCraft provides an open, community-driven model repository and a standard model format that allows prosumers to share and reuse those pretrained models. Actors can leverage ModelCraft's libraries to apply transfer learning and fine-tune the models using their data and expertise to fit their context-specific needs. Transfer learning allows only using a fraction of the training data and computational workload typically required. In contrast to sharing datasets, pretrained models do not require actual data sharing, alleviating possible data-sharing concerns. Stakeholders of ModelCraft share how the platform facilitates the sharing and refinement of models and the perceived benefits:

> *"So [the model repository] is a vehicle for academic teams to publish their models without sharing the data [... It is] for others to start with a pretrained model to accelerate the development process [...] I don't know, for a startup for a company, for another academic center, and fine-tune it on their data. [...] Whatever goes there, I would say the idea is not that this is a certified model that fulfills certain accuracy requirements that you can take and put in your product. Absolutely not. The idea is to share interesting models and allow the community to move faster. Because think what happened with the ImageNet pretrained backbones. It just speeds up everything."*
>
> (AI engineer, ModelCraft)

> *"On the transfer learning side, it's been pretty positive. We have proof points where you can, with about a fifth of the data, get your model up to the same accuracy. What he had was an example where at (site 1) they built a model, they sent it to (site 2) [...] and they were able to add a few of those [previously unseen] cases to the model, and then the model performed to the accuracy (site 1) had seen."*
>
> (Product manager, ModelCraft)

In conclusion, AI platforms apply the logic of distributed refinement to cultivate increasingly specialised datasets and models available for AI development on the platform. Through continuously refining resources by third parties, AI platforms can support AI development for a broader range of industry-specific use cases.

4.4 | Application brokering

In the orchestration logic of application brokering, AI platforms facilitate the transactions between distinct sides of producers and consumers of readily usable AI applications. Specifically, platforms like ClinDeploy and MedConnect, which focus on the model deployment and verification phases, facilitate the search and matchmaking of these applications. The producing side includes vendors of AI applications seeking to access the platform's user base and commercialise their products. The consuming side includes end-users such as hospitals and radiology centers seeking to access readily trained applications that fit their individual context and data environment.

Through this orchestration logic, AI platforms facilitate the transactions of AI applications in the market through matchmaking and reducing transaction costs. AI platforms aim to achieve this through curating third-party applications (e.g., verifying their generalizability), creating transparency of the applications available (e.g., on a marketplace), and proposing and enforcing standardised formats and interfaces (e.g., defining model specifications). A key challenge lies in estimating the true performance of external applications on local data, which can vary significantly in practice. Therefore, AI platforms make those applications comparable and triable, allowing consumers to benchmark competing products and make informed decisions.

For example, platforms like MedConnect provide an operating system and curated marketplace with over 50 external applications. MedConnect engages in a detailed analysis to review and preselect the applications offered on the platform (e.g., regulatory clearance). Consumers only need to contract and integrate with the platform once, thus considerably reducing transaction costs. Furthermore, MedConnect enables consumers to try and compare different applications to see how they perform on their local data before purchasing. Producers benefit from the installed user base, as the platform provides trusted access to potential customers. Stakeholders from MedConnect report how this orchestration logic facilitates transactions:

> *"The idea behind a platform is that you can reduce a lot of that overhead by just contracting once for the platform […] and then once you have that platform in place, you'll have access to all these different applications. […] Apart from the resources saved on the procurement aspects, it's also the value behind the integration and natural deployment, so typically what's called the last mile challenges."*
>
> (Clinical lead, MedConnect)

> *"We have a curated marketplace, which means really selected applications and not every lung nodule detection module that is available on the market. […] We have physicians in our team who look at the study results and talk to the manufacturers. What speaks for the evidence, how does the whole thing work? […] Medical evidence is one thing, but also requirements for data protection, data security, etc., in other words, that it really works in the relevant markets."*
>
> (Business manager, MedConnect)

> *"[We] harness [the customer's] data, make sure it's anonymized and ready for trial purposes, process that through any given algorithm, and present the results back to [the customer]. […] Here's how it works on your data to then allow that customer to have a very level playing field assessment of each one of those applications. What has become very apparent is that all the algorithms that are out in the market work very well. But you can see massive changes in their performance as you move them in different geographical territories across the world."*
>
> (Director of partnerships, MedConnect)

In conclusion, AI platforms apply the logic of application brokering to address the high burdens and uncertainties involved in transacting context-sensitive AI applications, enabling consumers to select applications that match their local data needs.

5 | DISCUSSION

Industry-specific AI platforms are confronted with distinct orchestration challenges, including the need for specialised resources, data-sharing concerns, heterogeneous user requirements, and context-sensitive AI applications (e.g., Hosny et al., 2018; Thrall et al., 2018). The failure to address these orchestration challenges hinders the scalability and effectiveness of industry-specific AI platforms, ultimately resulting in structural innovation and market failures (Jacobides et al., 2024).

The results reveal four orchestration logics that can build on each other to address these challenges. The first logic, *platform resourcing*, provides resources to individually curate data, train models and develop applications, mobilising the specialised data and expertise of otherwise resource-constrained actors. *Data-centric collaboration* builds on this logic by facilitating trusted collaboration, where dispersed data can be combined into diverse datasets and generalizable models, alleviating data-sharing concerns and overcoming heterogeneous user requirements. The third logic, *distributed refinement*, primarily enables co-specialised actors to share and refine datasets and models, producing increasingly specialised resources. This process yields context-specific applications ready for consumption or distribution. Building on these logics, *application brokering* matches competing third-party applications with consumers, who benchmark the performance based on their specific data and context.

Figure 3 illustrates this interplay of orchestration logics, highlighting how the platform coordinates autonomous actors' efforts and resource contributions to pass certain thresholds (e.g., generalizability of models). In the following, we discuss more thoroughly how the logics can address the challenges of specialised resources, data-sharing concerns, heterogeneous user requirements, and context-sensitive applications (see Table 4). We further elaborate on the generalizability of the orchestration logics to other industries, such as finance, healthcare, and manufacturing.

First, providing access to specialised resources (e.g., training data) to facilitate innovation in industry settings can be challenging and presents a threshold for specialised resource availability (Haki et al., 2024). Therefore, the logic of platform resourcing describes how platform owners provide software, developer, and infrastructure services such as developer tools and AI frameworks (Jacobides et al., 2021; Lins et al., 2021). Actors with access to unique data and specialised expertise use these services to mobilise their resources and develop datasets, models and applications (Geske et al., 2021). We see a similar logic in enterprise systems, where platform owners offer APIs and SDKs,

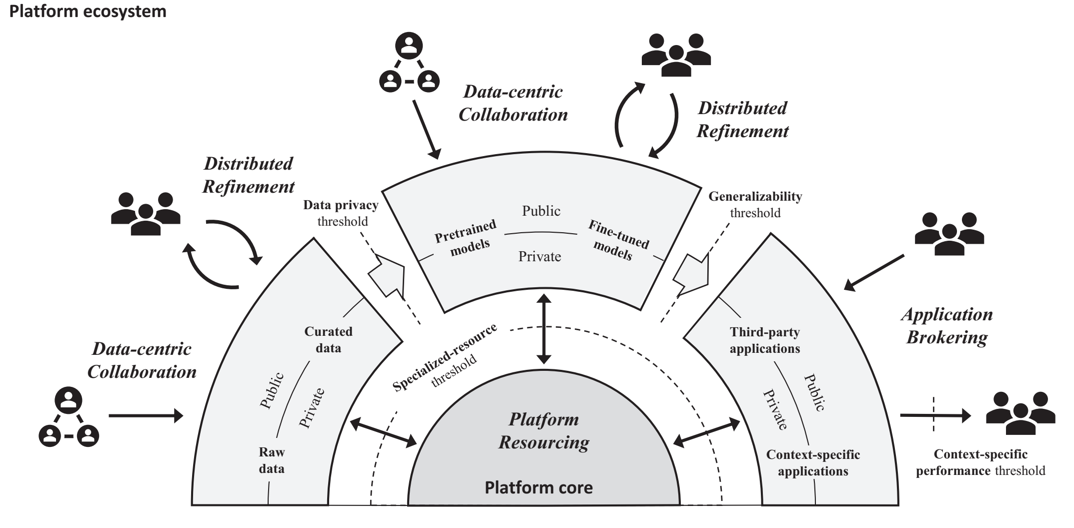

FIGURE 3 Four orchestration logics of industry-specific artificial intelligence platforms.

TABLE 4 Addressing orchestration challenges on industry-specific AI platforms.

Orchestration challenge	Platform orchestration logics
Providing access to highly specialised resources	Platform resourcing simplifies the initial creation of specialised resources for AI development, allowing otherwise constrained actors to mobilise their resources (e.g., data and domain expertise). Distributed refinement allows co-specialised actors to share and refine the resources in a distributed manner, stimulating innovation and providing increasingly specialised resources (e.g., pretrained models).
Data-sharing concerns	Data-centric collaboration facilitates the trusted integration of data resources from multiple actors (e.g., through federated learning). Distributed refinement allows actors to share and reuse models instead of actual data to alleviate data-sharing concerns (e.g., using transfer learning).
Heterogeneity of user requirements	Platform resourcing offers standardised interfaces for applications to integrate with users' heterogeneous up- and downstream systems. Data-centric collaboration facilitates the creation of more diverse datasets, yielding models that better generalise over heterogeneous users.
Context-sensitive AI applications	Distributed refinement facilitates the cost-efficient forking of context-specific applications, building on specialised datasets and models. Application brokering allows consumers to benchmark and select from competing third-party applications using their local data.

Abbreviation: AI, artificial intelligence.

enabling actors to develop specialised applications that meet their industry-specific requirements (Ceccagnoli et al., 2012).

This logic can be extended through distributed refinement, where platform owners use repositories, data management tools, and approaches such as transfer learning (Pan & Yang, 2009) to enable prosumers to share and refine datasets and models. These shared resources act as boundary resources provided by the ecosystem, going beyond the traditional wisdom that the platform owner provides these resources and then tuned by the ecosystem (Eaton et al., 2015). In an ongoing process, other prosumers refine these resources using their unique expertise and data, resulting in highly curated datasets and fine-tuned models. These specialised resources can cater to a broad range of industry-specific use cases, passing the threshold for specialised resource availability (Haki et al., 2024).

Second, data-sharing concerns present ongoing challenges in many industries, such as finance, healthcare and manufacturing (Jussen et al., 2024; Li et al., 2021). The findings show how—building on platform resourcing—data-centric collaboration allows actors to become trusted peer collaborators that publicly or privately integrate data from multiple sites. For example, in healthcare, several actors join forces to co-create open datasets for various occasions (Willemink et al., 2020). In other industries, such as manufacturing, we see data spaces as more formalised, decentralised infrastructures that govern the trusted data-sharing between autonomous actors (Möller et al., 2024; Otto & Jarke, 2019).

Alternatively, AI platforms can leverage federated learning as a computationally driven, privacy-preserving approach to combine peer collaborators' data (Agahari et al., 2022). In federated learning, the platform owner orchestrates actors to collaborate on training models without publicly sharing the curated data (Bi et al., 2023; Li et al., 2021). This approach is more suited for industries with stringent data privacy regulations. For example, financial institutions can use federated learning to develop models for loan prediction, training a model on their local data while only sharing model updates in the federation (Li et al., 2021).

Platform owners can further address data-sharing concerns through the logic of distributed refinement. Instead of sharing actual data, prosumers can share models as "processed data," passing the threshold of data privacy concerns. To enable this decoupling of models, platform owners provide model repositories, transfer learning tools, and federated learning frameworks (Q. Li et al., 2021). Reusing external models addresses the need for large datasets while reducing the required computing power, two significant constraints in AI development (Jacobides et al., 2021;

Jöhnk et al., 2021). However, building on external models may not offer the same flexibility as working with actual datasets, indicating a possible trade-off (Pan & Yang, 2009).

Third, the large heterogeneity of user requirements limits the scope of applications, reducing incentives for third-party producers to innovate and extend the platform (Rietveld & Eggers, 2018). To that end, platform resourcing can offer standardised interfaces with seamless integration to users' up- and downstream business systems. For example, industrial platforms can provide standardised integration with devices and production equipment to develop applications in manufacturing (Hein et al., 2019; Pauli et al., 2021). Through these standardised boundary resources, AI platforms can facilitate third-party innovation despite heterogeneous markets (Haki et al., 2024).

Beyond these integration issues, producers in AI development face heterogeneous data environments across users, which must be reflected in the training data (Brecker et al., 2023; Diaferia et al., 2022). For example, these data environments comprise different patient populations in healthcare (Hosny et al., 2018; Willemink et al., 2020) or different production setups in manufacturing (Weber et al., 2022). Therefore, AI platforms engage in data-centric collaboration to guide the co-creation of more extensive and diverse datasets. Sharing these datasets publicly or privately enables producers to create applications that can better generalise over heterogeneous user environments, allowing them to pass the threshold of generalizability (e.g., Willemink et al., 2020).

Last, AI applications are highly context-sensitive based on the data they are trained on (Brecker et al., 2023; Diaferia et al., 2022), challenging the matchmaking of producers and consumers. The findings show that platform owners support the distributed refinement of resources using additional data and expertise to fork several context-specific applications. For example, prosumers can build on a pretrained model specialised for the safety monitoring of construction workers and fine-tune it to fit different production contexts (e.g., Weber et al., 2022). This approach reduces the effort required for AI development (Diaferia et al., 2022), fostering the cost-efficient production of context-specific applications.

Building on the previous logics, application brokering provides access to competing third-party applications, widening the list of potentially suitable applications for consumers. These applications present standardised products for industry-specific use cases, offering opportunities for specialisation in different contexts (e.g., applications for specific machines) (Boudreau, 2012). The platform owner curates these applications through governance mechanisms that ensure regulatory conformity and sufficient generalizability (e.g., on independent data). In addition, the platform facilitates benchmarking competing applications using local consumer data, enhancing transparency over their contextual performance (Brecker et al., 2023). Thereby, consumers only select applications that cross their context-specific performance threshold. This logic allows platform owners to match producers and consumers of AI applications (Jacobides et al., 2021), despite their context sensitivity.

6 | IMPLICATIONS, LIMITATIONS AND FUTURE RESEARCH

Our findings have theoretical implications for the literature on platform orchestration in general and AI platforms in particular. In addition, we illustrate the practical implications of our results for platform owners and platform users. Last, we present the limitations of this study and provide suggestions for future research.

6.1 | Theoretical implications

The literature on platform orchestration has primarily focused on homogeneous, less complex markets (De Reuver et al., 2018; Rietveld & Schilling, 2021), centering on two orchestration logics: fostering third-party innovation through boundary resources (Eaton et al., 2015; Ghazawneh & Henfridsson, 2013) and facilitating transactions in multi-sided markets (Rochet & Tirole, 2003; Shi, 2023). However, industry-specific AI platforms face more complex challenges that these traditional logics do not fully address. To fill this gap, we introduce two additional orchestration logics and illustrate their interplay in overcoming industry-specific orchestration challenges.

More specifically, our study complements the established logics of innovation and transaction (Cusumano et al., 2019; Evans & Gawer, 2016) by proposing data-centric collaboration and distributed refinement. These two new logics actively guide autonomous actors in co-creating and refining datasets and models for AI development. This stands in contrast to leaving such between-actor interactions to the self-organisation of the ecosystem (Engert et al., 2024; Haki et al., 2024). The interplay of these orchestration logics enables platform owners to manage specialised resources better, facilitate collaboration despite data-sharing concerns and ensure that applications can adapt to heterogeneous, context-sensitive environments. By identifying and elaborating on these additional logics, our study offers a more comprehensive framework for understanding platform orchestration in complex settings, such as industry-specific AI development.

Furthermore, the literature on AI platforms has focused on delineating different types of AI platforms and the boundary resources they provide, such as AI software, developer, and infrastructure services (Geske et al., 2021; Jacobides et al., 2021; Lins et al., 2021). Despite noting a novel data layer (Li & Kettinger, 2021), this literature focused on providing traditional boundary resources to support the development and consumption of AI applications (Geske et al., 2021; Jacobides et al., 2021). We extend this line of research by shifting the perspective of boundary resources toward a process of boundary processing enabled through the decoupling of data and models.

Extant literature conceptualises boundary resources like AI developer tools as being provided by platform owners (Ghazawneh & Henfridsson, 2013) and adjusted through distributed actions within the ecosystem (Eaton et al., 2015). However, in complex markets like industry-specific AI development, platform owners cannot provide all the necessary resources (Haki et al., 2024). Therefore, our findings show that platform owners do more than provide predefined resources—they orchestrate a dynamic, iterative process of boundary processing, combining four orchestration logics to enable actors to co-create and refine data and models at the boundary (cf. Figure 3). This ongoing refinement allows for the flexible adaptation of models to meet heterogeneous user requirements and contextual settings. By shifting the perspective from platform-provided boundary resources to ecosystem-driven boundary processing, we offer a refined understanding of how data and models function as boundary resources in AI platforms (Geske et al., 2021; Jacobides et al., 2021; Rai et al., 2019).

Boundary processing requires decoupling data and model resources from the platform, which platform owners should consider in their architecture design. This finding extends the traditional view of decoupling for platforms, where the platform owner controls the infrastructure and decouples it from the applications provided by third parties (Benlian et al., 2018; Tiwana et al., 2010). Within AI platforms, we see that platform owners additionally decouple data and model resources to enable collaboration and innovation on these layers. Platform owners can actively enable this decoupling by providing repositories, data management tools, transfer learning, and federated learning (Li et al., 2021; Pan & Yang, 2009). However, the decoupling of models yields trade-offs: while actors can interact through models to alleviate data-sharing concerns and computing demands, these models offer less flexibility for innovation (Pan & Yang, 2009). These findings suggest that the decoupling of resources is more complex in AI platforms and requires platform owners to balance inclusiveness and flexibility in industry-specific AI development.

6.2 | Practical implications

The findings also have practical implications for platform owners and platform users. First, platform owners should proactively address the orchestration challenges unique to their industry, such as the need for specialised resources or strict data-sharing constraints. Our study highlights how the strategic use of orchestration logics can help overcome these barriers (see Table 4). For instance, in manufacturing, AI platforms can mobilise highly curated datasets for predictive maintenance by facilitating distributed refinement and data sharing. Trusted consortia are already advancing this by sharing specialised data (Möller et al., 2024), and expanding these efforts to include pretrained models could further stimulate innovation. In sectors like finance, where data-sharing is regulated, privacy-preserving mechanisms like federated learning can enable data-centric collaboration without compromising privacy (Bi et al., 2023; Li et al., 2021). Platform owners integrating these strategies into their platform design can effectively orchestrate AI development and drive innovation in industry-specific contexts.

Second, actors should explore deeper engagement with AI platforms beyond passive service consumption. Distributed refinement, for instance, demonstrates the value of contributing datasets or models to the ecosystem, enabling co-specialised actors to enhance and refine these resources. Although this may not yield immediate, direct returns, it creates long-term ecosystem benefits, such as improved models and applications for all participants. Similarly, engaging in data-centric collaboration with other actors enhances the generalizability and robustness of models (Thrall et al., 2018), offering benefits that isolated service consumption cannot achieve. Active participation thus unlocks the collective potential of the platform and its ecosystem.

Third, producers of AI applications must weigh the advantages and challenges of joining industry-specific platforms. Data-centric collaboration allows producers to leverage diverse datasets from heterogeneous users, resulting in more robust and generalizable applications. However, this collaboration introduces complexities, such as managing ongoing partnerships and incentivising cooperation (Bi et al., 2023). Application brokering offers access to an established consumer base, which is particularly valuable for smaller vendors and startups, but it also intensifies competition by enabling direct benchmarking of competing applications (Boudreau, 2012). Careful assessment of these trade-offs can help producers make informed decisions about participating in industry-specific AI platforms.

6.3 | Limitations and future research

Our study is not free from limitations regarding its internal and external validity. Regarding internal validity, our case study research approach is inherently tied to subjective interpretations and researcher bias. We aimed to address

TABLE 5 Future research areas and exemplary questions.

Research area	Exemplary research questions
Platform orchestration for industry-specific AI platforms	• What are the long-term implications of the orchestration logics, such as the impact of data-centric collaboration on application quality? • How can more nuanced orchestration logics inform other types of data-driven platforms, such as industrial platforms (Pauli et al., 2021)?
Data and AI models as boundary resources	• Which characteristics enhance or restrict the generative potential of data and models as boundary resources? • Compared to data, how do models as privacy-preserving and computationally efficient boundary resources impact collaboration and innovation? • How can federated data concepts, such as data spaces (Möller et al., 2024), support using data as boundary resources on AI platforms?
Platform governance for industry-specific AI platforms	• How can industry-specific AI platforms sustain third- party application contributions despite intensified competition and individual benchmarking (Engert et al., 2023)? • Given constantly changing data, how can platform owners effectively govern the long-term performance and evolution of applications (Paleyes et al., 2022)? • How should platform owners govern boundary processing of data and models, and should those two resources be governed differently?
Platform competition of industry-specific AI platforms	• How do industry-specific AI platforms co-evolve with other AI platforms and key players (e.g., AWS, Google AI) and adapt to the fast progress of AI technology? • Given the foundational role of data contributions, do industry-specific AI platforms also inhibit data-driven network effects based on their installed base (Gregory et al., 2021)?

Abbreviation: AI, artificial intelligence.

researcher bias by following structured data analysis techniques (Miles et al., 2018), constantly comparing our emergent findings with literature (Strauss & Corbin, 1990), and critically reflecting our presumptions within the research team (Walsham, 2006). Furthermore, we aimed to increase internal validity by capturing the views of multiple platform stakeholders in our interviews and by grounding our findings in five empirical cases of AI platforms to enhance the robustness of our findings (Eisenhardt, 1989).

Regarding external validity, our study is limited to applications that build on deep learning techniques, specifically supervised learning (Russell & Norvig, 2021). Other AI techniques, such as generative AI, may be subject to other challenges and potentially require different approaches to platform orchestration. Furthermore, our study draws on insights from the medical imaging sector, which has unique characteristics, such as the high sensitivity of patient data (Hosny et al., 2018; Thrall et al., 2018). However, we are confident that the derived orchestration logics are abstract enough to generalise to other industry contexts. For example, as discussed earlier, the finance and manufacturing industries face similar challenges, such as the need for specialised resources and pronounced data-sharing concerns (e.g., Bi et al., 2023; Li et al., 2021).

Beyond addressing these limitations, the implications of our study point to additional future research avenues. First, there is a need to further investigate data and models as distinct types of resources processed at the boundary. Second, we propose to examine effective platform governance for industry-specific AI platforms, given the unique orchestration logics and challenges identified in this study (see also Li and Kettinger (2021)). Last, future research could examine how AI platforms compete and adapt over time, given their important role in orchestrating industry-specific AI development and the dynamic evolution of the AI ecosystem (Jacobides et al., 2021). We provide exemplary future research questions in Table 5.

7 | CONCLUSION

Industry-specific AI platforms face orchestration challenges related to specialised resources, data-sharing concerns, heterogeneous user requirements, and context-sensitive applications. Our theoretical understanding has been limited in addressing these challenges due to a research focus on more homogeneous and less complex markets. Our study of five AI platforms in medical imaging reveals four orchestration logics to address these challenges: platform resourcing, data-centric collaboration, distributed refinement and application brokering. These logics illustrate how platform owners can verticalize the AI development process by orchestrating autonomous actors who co-create, share and refine data and models, eventually producing context-specific applications capable of generalisation. The resulting applications can be offered as standardised products but require competitive benchmarking by consumers.

As AI technology evolves and its applications become increasingly tailored to industry-specific needs, AI platforms will likely play a key role in orchestrating the journey from raw data to industry-specific applications. This study's insights advance our theoretical understanding of platform orchestration for AI platforms and offer practical guidance to platform owners and platform users. Effectively orchestrating complex ecosystems will be essential to realising AI platforms' full potential and creating value within specific industries.

ACKNOWLEDGMENTS

The authors want to thank the senior editor, associate editor, three anonymous reviewers, and the 2022 ISJ Trending Topics Workshop participants for their helpful and constructive feedback. We further want to thank Tobias Jenne and Marc Sommer, who contributed to this research project's data collection and analysis. We are further indebted to the practitioners who shared valuable insights for this research, including esteemed experts from the American College of Radiology. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

Author elects to not share data.

ORCID

Michael Weber https://orcid.org/0000-0003-1528-1856
Jörg Weking https://orcid.org/0000-0002-5288-240X

REFERENCES

Agahari, W., Ofe, H., & de Reuver, M. (2022). It is not (only) about privacy: How multi- party computation redefines control, trust, and risk in data sharing. *Electronic Markets*, *32*(3), 1577–1602.

Ashmore, R., Calinescu, R., & Paterson, C. (2021). Assuring the machine learning lifecycle: Desiderata, methods, and challenges. *ACM Computing Surveys (CSUR)*, *54*(5), 1–39.

Baldwin, C., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, markets and innovation*. Edward Elgar.

Benlian, A., Kettinger, W. J., Sunyaev, A., & Winkler, T. J. (2018). The transformative value of cloud computing: A decoupling, platformization, and recombination theoretical framework. *Journal of Management Information Systems*, *35*(3), 719–739.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *Mis Quarterly*, *45*(3), 1433–1450.

Bi, X., Gupta, A., & Yang, M. (2023). Understanding partnership formation and repeated contributions in federated learning: An analytical investigation. *Management Science*, *70*(8), 4974–4994.

Boudreau, K. (2010). Open platform strategies and innovation: Granting access vs. devolving control. *Management Science*, *56*(10), 1849–1872.

Boudreau, K. (2012). Let a thousand flowers bloom? An early look at large numbers of software app developers and patterns of innovation. *Organization Science*, *23*(5), 1409–1427.

Brecker, K., Lins, S., Trenz, M., & Sunyaev, A. (2023). Artificial intelligence as a service: Trade-offs impacting service design and selection. In *ICIS 2023 Proceedings*. 5, Hyderabad, India. https://aisel.aisnet.org/icis2023/itadopt/itadopt/5

Ceccagnoli, M., Forman, C., Huang, P., & Wu, D. (2012). Cocreation of value in a platform ecosystem! The case of enterprise software. *MIS Quarterly*, *36*, 263–290.

Cohen, I. G., & Mello, M. M. (2019). Big data, big tech, and protecting patient privacy. *JAMA*, *322*(12), 1141–1142.

Cusumano, M., Gawer, A., & Yoffie, D. B. (2019). *The business of platforms: Strategy in the age of digital competition, innovation, and power*. Harper Business.

De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: A research agenda. *Journal of Information Technology*, *33*(2), 124–135.

Diaferia, L., Blohm, I., De Rossi, L. M., & Salviotti, G. (2022). When standard is not enough: A conceptualization of AI Systems' customization and its antecedents. In *ICIS 2022 Proceedings*. 8, Copenhagen, Denmark. https://aisel.aisnet.org/icis2022/is_implement/is_implement/8

Duda, S., Hofmann, P., Urbach, N., Völter, F., & Zwickel, A. (2023). The impact of resource allocation on the machine learning lifecycle. *Business & Information Systems Engineering*, *66*(2), 203–219.

Eaton, B., Elaluf-Calderwood, S., Sørensen, C., & Yoo, Y. (2015). Distributed tuning of boundary resources. *MIS Quarterly*, *39*(1), 217–244.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, *50*(1), 25–32.

Eisenmann, T., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard Business Review*, *84*(10), 92.

Engert, M., Evers, J., Hein, A., & Krcmar, H. (2023). Sustaining complementor engagement in digital platform ecosystems: Antecedents, behaviours and engagement trajectories. *Information Systems Journal*, *33*(5), 1151–1185.

Engert, M., Hein, A., Maruping, L. M., Thatcher, J., & Krcmar, H. (2024). Self-organization and governance in digital platform ecosystems: An information ecology approach. *Mis Quarterly*, forthcoming. https://doi.org/10.25300/MISQ/2024/18413

Evans, P. C., & Gawer, A. (2016). *The rise of the platform enterprise: A global survey*. The Emerging Platform Economy Series.

Foerderer, J., Kude, T., Schuetz, S. W., & Heinzl, A. (2019). Knowledge boundaries in enterprise software platform development: Antecedents and consequences for platform governance. *Information Systems Journal*, *29*(1), 119–144.

Geske, F., Hofmann, P., Lämmermann, L., Schlatt, V., & Urbach, N. (2021). Gateways to Artificial Intelligence: Developing a Taxonomy for AI Service Platforms. Paper presented at the 29th European conference on information systems (ECIS), virtual conference.

Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. *Information Systems Journal*, *23*(2), 173–192.

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods*, *16*(1), 15–31.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2021). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review*, *46*(1), 534–551.
Haki, K., Tanriverdi, H., Safaei, D., Schmid, M., Aier, S., & Winter, R. (2024). Generativity and profitability on B2B innovation platforms: A simulation- based theory development. *MIS Quarterly*, *48*(2), 583–612.
Hein, A., Weking, J., Schreieck, M., Wiesche, M., Böhm, M., & Krcmar, H. (2019). Value co-creation practices in business-to-business platform ecosystems. *Electronic Markets*, *29*(3), 503–518.
Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H., & Aerts, H. J. (2018). Artificial intelligence in radiology. *Nature Reviews Cancer*, *18*(8), 500–510.
Jacobides, M. G., Brusoni, S., & Candelon, F. (2021). The evolutionary dynamics of the artificial intelligence ecosystem. *Strategy Science*, *6*(4), 412–435.
Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, *39*(8), 2255–2276.
Jacobides, M. G., Cennamo, C., & Gawer, A. (2024). Externalities and complementarities in platforms and ecosystems: From structural solutions to endogenous failures. *Research Policy*, *53*(1), 104906.
Jöhnk, J., Weißert, M., & Wyrtki, K. (2021). Ready or not, AI comes—An interview study of organizational AI readiness factors. *Business & Information Systems Engineering*, *63*(1), 5–20.
Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.
Jussen, I., Möller, F., Schweihoff, J., Gieß, A., Giussani, G., & Otto, B. (2024). Issues in inter-organizational data sharing: Findings from practice and research challenges. *Data & Knowledge Engineering*, *150*, 102280.
Li, H., & Kettinger, W. B. J. (2021). The building blocks of software platforms: Understanding the past to forge the future. *Journal of the Association for Information Systems*, *22*(6), 1524–1555.
Li, Q., Wen, Z., Wu, Z., Hu, S., Wang, N., Li, Y., … He, B. (2021). A survey on federated learning systems: Vision, hype and reality for data privacy and protection. *IEEE Transactions on Knowledge and Data Engineering*, *35*(4), 3347–3366.
Lins, S., Pandl, K. D., Teigeler, H., Thiebes, S., Bayer, C., & Sunyaev, A. (2021). Artificial intelligence as a service. *Business & Information Systems Engineering*, *63*(4), 441–456.
Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook* (4th ed.). Sage publications.
Möller, F., Jussen, I., Springer, V., Gieß, A., Schweihoff, J. C., Gelhaar, J., … Otto, B. (2024). Industrial data ecosystems and data spaces. *Electronic Markets*, *34*(1), 1–17.
Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, *17*(1), 2–26.
Otto, B., & Jarke, M. (2019). Designing a multi-sided data platform: Findings from the international data spaces case. *Electronic Markets*, *29*(4), 561–580.
Paleyes, A., Urma, R.-G., & Lawrence, N. D. (2022). Challenges in deploying machine learning: A survey of case studies. *ACM Computing Surveys (CSUR)*, *55*(6), 1–26.
Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359.
Pauli, T., Fielt, E., & Matzner, M. (2021). Digital industrial platforms. *Business & Information Systems Engineering*, *63*(2), 181–190.
Pumplun, L., Tauchert, C., & Heidt, M. (2019). A new organizational chassis for artificial intelligence-exploring organizational readiness factors. In *Proceedings of the 27th European conference on information systems (ECIS)*, Stockholm & Uppsala, Sweden. https://aisel.aisnet.org/ecis2019_rp/106/
Rai, A., Constantinides, P., & Sarker, S. (2019). Editor's comments: Next-generation digital platforms: Toward human–AI hybrids. *MIS Quarterly*, *43*(1), iii–x.
Rajpurkar, P., Chen, E., Banerjee, O., & Topol, E. J. (2022). AI in health and medicine. *Nature Medicine*, *28*(1), 31–38.
Rietveld, J., & Eggers, J. (2018). Demand heterogeneity in platform markets: Implications for complementors. *Organization Science*, *29*(2), 304–322.
Rietveld, J., & Schilling, M. A. (2021). Platform competition: A systematic and interdisciplinary review of the literature. *Journal of Management*, *47*(6), 1528–1563.
Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, *1*(4), 990–1029.
Roth, H. R., Cheng, Y., Wen, Y., Yang, I., Xu, Z., Hsieh, Y.-T., … Lu, K. (2022). Nvidia flare: Federated learning from simulation to real-world.
Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
Shi, P. (2023). Optimal matchmaking strategy in two-sided marketplaces. *Management Science*, *69*(3), 1323–1340.
Siggelkow, N. (2007). Persuasion with case studies. *Academy of Management Journal*, *50*(1), 20–24.
Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., … Teller, A. (2016). Artificial intelligence and life in 2030. In *One hundred year study on artificial intelligence (report of the 2015–2016 study panel)*. Stanford. http://ai100.stanford.edu/2016-report

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Sage Publications, Inc.

Thrall, J. H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., & Brink, J. (2018). Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. *Journal of the American College of Radiology*, *15*(3), 504–508.

Tiwana, A. (2014). The rise of platform ecosystems. In *Platform ecosystems: Aligning architecture, governance, and strategy*. Morgan Kaufmann.

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Research commentary—Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research*, *21*(4), 675–687.

van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the machine meets the expert: An ethnography of developing AI for hiring. *MIS Quarterly*, *45*(3), 1557–1580.

van Leeuwen, K. G., Schalekamp, S., Rutten, M. J. C. M., van Ginneken, B., & de Rooij, M. (2021). Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *European Radiology*, *31*(6), 3797–3804.

Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems*, *15*(3), 320–330.

Wareham, J., Fox, P. B., & Cano Giner, J. L. (2014). Technology ecosystem governance. *Organization Science*, *25*(4), 1195–1215.

Weber, M., Engert, M., Schaffer, N., Weking, J., & Krcmar, H. (2023). Organizational capabilities for AI implementation—Coping with inscrutability and data dependency in AI. *Information Systems Frontiers*, *25*, 1549–1569.

Weber, M., Pfeiler, M., Hein, A., Weking, J., & Krcmar, H. (2022). Deploying AI applications to multiple environments: Coping with environmental, data, and predictive variety. In *ICIS 2022 Proceedings*. 7, Copenhagen, Denmark. https://aisel.aisnet.org/icis2022/is_design/is_design/7/

Willemink, M. J., Koszek, W. A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., … Lungren, M. P. (2020). Preparing medical imaging data for machine learning. *Radiology*, *295*(1), 4–15.

Wlcek, M., Knapp, J., Marheine, C., & Wortmann, F. (2023). Revisiting the IIoT platform graveyard: Key learnings from failed IIoT platform initiatives. In *ICIS 2023 Proceedings*. 6, Hyderabad, India. https://aisel.aisnet.org/icis2023/practitioner/practitioner/6

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE.

AUTHOR BIOGRAPHIES

Michael Weber is a Research Associate and doctoral student at the TUM School for Computation, Information and Technology of the Technical University of Munich (TUM), Germany. In his research, Michael explores the wider implications of artificial intelligence technology on firms, including its impact on business models and inter-organisational collaboration. His research has appeared in Business & Information Systems Engineering, Information Systems Frontiers and refereed conference proceedings such as ICIS and ECIS.

Andreas Hein is an Assistant Professor of IT Management affiliated with the Institute of Information Systems and Digital Business in the School of Management of the University of St. Gallen, Switzerland. Before that, he was a postdoctoral researcher at the TUM School for Computation, Information and Technology of the Technical University of Munich (TUM), Germany, where he also received his PhD. In addition, Andreas has 3 years of experience as a Senior Strategy Consultant at IBM. His work has appeared in the MIS Quarterly, the Journal of Strategic Information Systems, the Information Systems Journal, the European Journal of Information Systems and the Government Information Quarterly, as well as refereed conference proceedings such as ICIS and ECIS. Andreas focuses his research on IT management at the organisational level, digital platform ecosystems and eGovernment.

Jörg Weking is a Technical Specialist in artificial intelligence at IBM. Before that, he was a postdoctoral researcher at the TUM School for Computation, Information and Technology of the Technical University of Munich (TUM), Germany, where he also received his doctoral degree. He has been a visiting fellow at the Queensland University of Technology (QUT), Australia. Jörg's research interests include digital business models, digital transformation, and artificial intelligence. His work has appeared in the Journal of Strategic Information Systems, the Information Systems Journal, the European Journal of Information Systems, and the International Journal of Production Economics, as well as refereed conference proceedings such as ICIS and ECIS.

Helmut Krcmar is a Professor Emeritus of Information Systems and "TUM Emeritus of Excellence" at the Technical University of Munich (TUM), Germany. He received his doctoral degree from the Saarland University, Germany, and an honorary doctoral degree from the University of St. Gallen, Switzerland. His research interests include digital transformation, platform-based ecosystems, information and knowledge management, service management, and information systems for healthcare and government. He is a Senior Scholar and Fellow of the Association for Information Systems (AIS). His work has appeared in the MIS Quarterly, Journal of Management Information Systems, Journal of Strategic Information Systems, Information Systems Journal, and MIS Quarterly Executive, among others.

How to cite this article: Weber, M., Hein, A., Weking, J., & Krcmar, H. (2024). Orchestration logics for artificial intelligence platforms: From raw data to industry-specific applications. *Information Systems Journal*, 1–29. https://doi.org/10.1111/isj.12567

APPENDIX A

TABLE A1 List of interviews.

ID	Case	Perspective	Role description	Round 1 (min)	Round 2 (min)
1	DataForge	Platform	Medical Advisor	57	
2	DataForge	Platform	Project Manager	28	
3	DataForge	Platform	Platform AI Engineer	43	
4	DataForge	Platform	Business Manager		29
5	DataForge	User	AI Engineer and Radiologist	41	
6	DataForge	User	Professor of Radiology	41	
7	DataForge	User	Radiologist and Data Annotator		33
8	RadiaHub	Platform	Vice President Data Science	23	
9	RadiaHub	Platform	Platform AI Engineer	33	
10	RadiaHub	Platform	Director of Platform	51	49*
11	RadiaHub	Platform	Chief Medical Information Officer		49*
12	RadiaHub	User	Medical Director of Radiology Informatics	32	
13	RadiaHub	User	AI Engineer and Researcher	35	
14	ModelCraft	Platform	Product Manager		39
15	ModelCraft	Platform	Account Manager		49
16	ModelCraft	Platform	AI Engineer		73
17	ModelCraft	Platform	Product Manager and Radiologist		38
18	ModelCraft	User	Radiologist and Project Manager		66
19	ModelCraft	User	Researcher, Radiologist, and Entrepreneur		39
20	ModelCraft	User	Medical Director		31
21	ClinDeploy	Platform	Chief Executive Officer	47	
22	ClinDeploy	Platform	Chief Partnerships Officer	36	
23	ClinDeploy	Platform	Chief Technology Officer	28	
24	ClinDeploy	Platform	Chief Operations Officer	38	
25	ClinDeploy	Platform	Business Development		34
26	ClinDeploy	User	Radiologist and Researcher	26	32
27	ClinDeploy	User	Professor for Radiology	23	
28	ClinDeploy	User	Senior Radiologist	15	
29	MedConnect	Platform	Head of Medical Affairs		36
30	MedConnect	Platform	Vice President Partnerships		38
31	MedConnect	Platform	Business Manager		36
32	MedConnect	Platform	Developer Program Lead		32
33	MedConnect	Platform	Clinical Lead		29
34	MedConnect	User	Chief Commercial Officer		31
35	MedConnect	User	AI Engineer and Product Manager		27
36	MedConnect	User	Chief Innovation Officer		30
37	MedConnect	User	Sales Manager		32
			Total no. of interviews	17	21

*Interview with both interviewees.

TABLE A2 Interview guidelines for the first round of data collection.

Interview guidelines for the first round of data collection
General questions (everyone): • Major barriers to AI development • General strategies to cope with development barriers • Own participation in AI development (e.g., exemplary projects)

Platform owner questions:
• Motives and offerings of the platform
    ○ Functionality of platform
    ○ Role of data on the platform
• Role and engagement of external actors

Platform user questions:
• Engagement with platform
    ○ Motives, expectations and alternatives
    ○ Experienced benefits and challenges

TABLE A3 Interview guidelines for the second round of data collection.

Interview guidelines for the second round of data collection
General questions (everyone): • Major barriers to AI development • Role of platforms to address barriers

Platform owner questions:
• Relevant partners and collaborators for the platform
• Supporting AI development
    ○ Detailed functioning of key offerings
    ○ Design challenges (e.g., integration, interoperability)
• Marketplaces/repositories for transactions
    ○ Functionality and expected value
    ○ Experiences (e.g., volume and perceived quality of transactions)
    ○ Incentive and control of third-party contributions
• Facilitating collaboration (e.g., federated learning, merging datasets)
    ○ Functionality and expected value (e.g., diversity of data)
    ○ Experiences (e.g., past projects)
    ○ Remaining challenges

Platform user questions:
• Other relevant partners and collaborators for AI development
• Experienced benefits and challenges with the platform (e.g., exemplary projects) for
    ○ Application development
    ○ Providing/consuming datasets and pretrained models
    ○ Providing/consuming applications (e.g., issues of transferability)
    ○ Collaborating with peers through the platform (e.g., the value of data diversity)

TABLE A4 Description of AI platform cases.

ID	Pseudonym	Case description
1	DataForge	DataForge is a commercial platform operated by a small firm in Northern America. DataForge strives to support AI research and development through high-quality data management. To that end, DataForge offers various services facilitating data preparation and annotation. Users can collaborate by sharing and combing datasets and collectively annotating data. In addition, DataForge allows users to share models to verify their performance on different datasets. DataForge's ecosystem comprises research institutions, hospitals, IT firms, AI vendors and annotators.
2	RadiaHub	RadiaHub is a non-profit platform operated by a research institute in Northern America. RadiaHub strives to democratise AI and support AI research and development. To that end, RadiaHub offers easy-to-use services that facilitate data management, model training and model verification. Users can collaborate by exchanging pretrained models, conducting federated learning and collaboratively working on datasets. RadiaHub further develops services for independent model verification and monitoring across different users. RadiaHub's ecosystem comprises research institutions, hospitals and IT firms.
3	ModelCraft	ModelCraft is an open-source platform backed by renowned IT firms and research institutions from Northern America and Europe. ModelCraft strives to support AI research and development and adopt AI applications in clinical practice. To that end, ModelCraft offers SDKs and libraries to support data management, model training, and model deployment. ModelCraft supports state-of-the-art concepts, such as active learning and federated learning, to drive the performance of AI applications. ModelCraft further hosts a repository for sharing and reusing pretrained models. ModelCraft's ecosystem comprises research institutions, hospitals, IT firms and AI vendors.
4	ClinDeploy	ClinDeploy is a commercial platform operated by a small-sized firm in Europe. ClinDeploy strives to support the adoption of AI applications in clinical practice. To that end, ClinDeploy offers an operating system and supporting services to facilitate application development, verification and deployment. ClinDeploy further provides a curated marketplace with third-party applications that users can benchmark and integrate into their clinical processes. ClinDeploy's ecosystem comprises hospitals, radiology service providers, research institutions and AI vendors.
5	MedConnect	MedConnect is a commercial platform operated by a large medical firm in North America and Europe. MedConnect strives to support AI research and development and adopt AI applications in clinical practice. To that end, MedConnect offers an operating system and supporting services to facilitate application development, verification and deployment. MedConnect further provides a curated marketplace with third-party applications. In addition, MedConnect offers selected actors access to their partner network as well as internal data and training infrastructure to support the development of new applications. MedConnect's ecosystem comprises hospitals, radiology service providers, research institutions, startups and AI vendors.