Manexia
Journal of Business, Management, and Creative Economy

Vol. 2 I Issue 1 I January 2026

Data Extractivism and Strategic Value Appropriation:
Rethinking Firm Advantage in AI-Centric SME Ecosystems

Taufik Wibisono[1]*

***Corresponding Mail:**
taufik.tik@bsi.ac.id

Article History:
Submitted: 29-08-2025
Approved: 08-11-2025
Published: 05-01-2026

OPEN ACCESS
Available at the open access journal:
https://sciedex.com/manexia

Abstrak

Artificial intelligence–enabled platforms are transforming the foundations of competitive advantage in digital market ecosystems. Small and medium-sized enterprises (SMEs) generate substantial transactional and behavioral data through platform participation, yet control over data aggregation and model-training architectures typically resides with platform sponsors. This structural decoupling challenges the core assumption of the resource-based view that ownership and control of valuable resources ensure rent appropriation. Integrating resource-based theory, value appropriation logic, data-enabled learning research, and platform governance scholarship, this article develops a conceptual framework explaining how data extractivism operates as an architecture-mediated mechanism of value capture. The model argues that competitive advantage in AI-centric ecosystems increasingly derives from control over aggregation infrastructures rather than localized data generation. Cross-SME data pooling produces compounding learning rents that disproportionately accrue to actors controlling centralized architectures, especially under conditions of high switching costs, limited data portability, and governance opacity. By reframing advantage as architecture-dependent, the study extends strategic management theory and clarifies how SME performance becomes ecosystem-conditioned in AI-driven markets.

Keywords

data extractivism; SME strategy; value appropriation; platform ecosystems; artificial intelligence governance; architecture-based competitive advantage

[1] Universitas Bina Sarana Informatika

1. Introduction

The rapid diffusion of artificial intelligence (AI) across digital platforms has reconfigured how value is created, distributed, and appropriated within contemporary market ecosystems. In platform-mediated environments, small and medium-sized enterprises (SMEs) increasingly rely on AI-driven infrastructures—recommendation systems, predictive analytics, dynamic pricing engines, and customer profiling tools—to access demand and enhance operational performance. While prior research has emphasized the productivity gains and dynamic capability enhancement associated with AI adoption (Raisch & Krakowski, 2021; Verhoef et al., 2021), less attention has been paid to a more structural question: who captures the value generated through AI-enabled data flows?

SMEs generate substantial volumes of transactional, behavioral, and contextual data through their participation in digital platforms. These data streams fuel machine learning models that optimize search, targeting, personalization, and pricing at scale. Yet the aggregation, processing, and monetization of cross-firm data are typically controlled by platform sponsors rather than by individual SMEs. As a result, the competitive advantage derived from data-enabled learning may accumulate disproportionately at the platform level, even when SMEs are the primary originators of raw data inputs. This structural asymmetry raises a theoretical tension within strategic management scholarship.

The resource-based view (RBV) posits that sustainable competitive advantage derives from firm-specific resources that are valuable, rare, inimitable, and non-substitutable (Barney, 1991). Central to this logic is the assumption that firms exercise effective control over the resources from which they derive rents. However, in AI-centric ecosystems, ownership, control, and appropriation of data-based resources are increasingly decoupled. SMEs may generate data, but platforms often control the architecture through which data are aggregated, recombined, and transformed into predictive capabilities. When control over data aggregation and model training resides outside the focal firm, the traditional alignment between resource ownership and rent appropriation becomes unstable.

Recent research has begun to recognize that digital platforms function as meta-organizations that coordinate and govern ecosystems of interdependent actors (Kretschmer et al., 2022). Governance mechanisms—such as access rules, API design, ranking algorithms, and monetization policies—shape the incentives and constraints faced by complementors (Chen et al., 2022). Yet this literature has primarily focused on coordination efficiency, innovation dynamics, and ecosystem growth. The distributional consequences of AI-enabled data aggregation for SME-level value capture remain under-theorized.

Parallel work in economics and strategy has demonstrated that data-enabled learning can generate compounding advantages over time (Hagiu & Wright, 2020). When machine learning systems improve as data accumulates, actors controlling larger and more diverse datasets may experience self-reinforcing performance gains. In multi-sided ecosystems, however, data are frequently pooled across heterogeneous complementors. Platforms may thus benefit from cross-SME data aggregation that no individual SME can replicate independently. This dynamic creates the potential for what may be termed learning rents—returns derived not merely from data possession but from control over the aggregation architecture that transforms distributed data into scalable intelligence.

The divergence between value creation and value appropriation has long been recognized in strategy research (Brandenburger & Stuart, 1996). Nonetheless, AI-centric ecosystems introduce a qualitatively different configuration. Rather than bargaining power alone determining surplus division, architectural control over data flows and learning systems becomes a central determinant of rent distribution. When platform sponsors design and control the infrastructure through which data are collected, processed, and monetized, appropriation becomes architecture-mediated rather than purely negotiation-based.

This architecture-mediated appropriation is particularly consequential for SMEs. Compared with large incumbents, SMEs typically possess limited data analytics capabilities, restricted access to cross-market information, and lower bargaining power within platform ecosystems. Although participation in digital platforms may enhance short-term visibility and sales performance, it may simultaneously deepen structural dependence on centralized AI infrastructures. Over time, such dependence may stabilize extractive equilibria in which SMEs continue contributing data while capturing a declining share of ecosystem-level rents.

Despite growing attention to platform power and digital governance (Gawer, 2021; Kretschmer et al., 2022), strategic management theory lacks an integrated account of how AI-driven data aggregation reshapes the logic of competitive advantage for SMEs. Existing AI strategy research has primarily examined organizational transformation, managerial cognition, and capability reconfiguration (Raisch & Krakowski, 2021). Meanwhile, ecosystem scholarship has emphasized complementor participation and innovation incentives without systematically theorizing surplus distribution under data asymmetry.

Accordingly, this study addresses the following research question:

How does data extractivism in AI-centric ecosystems reshape the mechanisms of value appropriation and competitive advantage for SMEs?

To answer this question, the article develops a mechanism-based conceptual framework integrating the resource-based view, value appropriation theory, and platform governance research. It advances three core arguments. First, in AI-centric ecosystems, sustainable advantage increasingly derives from control over data aggregation architectures rather than from data generation per se. Second, cross-actor data pooling generates compounding learning rents that may disproportionately benefit platform sponsors. Third, switching costs, data non-portability, and governance opacity can stabilize extractive equilibria that constrain SME-level value capture despite continued participation.

By reframing competitive advantage through the lens of architectural control and data extractivism, this study contributes to strategic management theory in three ways. It extends RBV by distinguishing between resource ownership and resource control in digitally mediated environments. It refines value appropriation theory by conceptualizing rent distribution as structurally embedded in platform design rather than solely determined by bargaining interactions. Finally, it situates SMEs at the center of AI strategy research, highlighting how ecosystem-level architectures condition firm-level performance outcomes.

In doing so, the article moves beyond narratives of digital empowerment and instead interrogates the structural conditions under which AI-enabled ecosystems redistribute value. Understanding these mechanisms is essential for advancing theory on competitive advantage in an era where data flows transcend firm boundaries and where architectural control increasingly shapes the distribution of economic rents.

2. Theoretical Foundations

The central theoretical challenge addressed in this article concerns the decoupling of value creation and value appropriation in AI-centric platform ecosystems. While SMEs generate economically valuable data through their market participation, control over data aggregation and AI learning systems often resides at the platform level. To theorize this structural asymmetry rigorously, this section integrates four streams of scholarship: (1) the resource-based view (RBV), (2) value creation and appropriation theory, (3) data-enabled learning and network effects, and (4) platform governance and ecosystem control. Together, these perspectives provide the analytical scaffolding for conceptualizing data extractivism as a structural feature of AI-driven ecosystems.

2.1 Resource-Based View and the Assumption of Resource Control

The resource-based view (RBV) posits that sustainable competitive advantage derives from resources that are valuable, rare, inimitable, and non-substitutable (Barney, 1991). Subsequent refinements have emphasized isolating mechanisms, causal ambiguity, and path dependence as sources of persistent rents (Peteraf, 1993). Implicit within this framework is a critical assumption: firms exercise effective control over the strategic resources from which they derive competitive advantage.

In traditional settings, ownership and control of key assets—whether physical capital, intellectual property, or organizational routines—were largely co-located within firm boundaries. However, digital transformation has progressively blurred these boundaries (Jacobides et al., 2018). In platform-mediated ecosystems, critical inputs into competitive advantage—such as user data, behavioral traces, and transactional histories—are generated across multiple actors and aggregated through centralized infrastructures.

Recent research has begun to reconsider the foundations of competitive advantage in digital environments. Hagiu and Wright (2020) demonstrate that data-enabled learning can generate dynamic advantages that intensify with scale, particularly when firms control the feedback loops linking user interactions to algorithmic improvement. Yet their models largely assume that the focal firm controls the data aggregation process. In AI-centric ecosystems, by contrast, SMEs may generate data while platforms control the architecture through which those data are transformed into predictive capabilities.

This decoupling challenges a core RBV premise. If SMEs generate valuable data but lack control over its aggregation, recombination, and monetization, then the link between resource generation and rent appropriation becomes unstable. Competitive advantage may shift from resource ownership to architectural control—the authority to design, aggregate, and operationalize distributed data assets. Such a shift necessitates revisiting RBV through a control-rights lens, emphasizing governance over data flows rather than mere possession of data inputs.

2.2 Value Creation and Value Appropriation in Interdependent Systems

Strategy research has long distinguished between value creation and value appropriation (Brandenburger & Stuart, 1996). Total value created within a system can exceed the share captured by any single participant; the division of surplus depends on bargaining power, outside options, and competitive positioning. In vertically integrated industries, surplus division often reflects negotiation between clearly defined actors. However, platform ecosystems introduce multi-sided interdependencies that complicate traditional bargaining models.

Jacobides et al. (2018) argue that ecosystems differ from industries in that complementarities and modular architectures structure inter-firm relationships. Within such systems, control over bottlenecks—points of interdependence where value flows converge—can shape rent distribution. In AI-centric platforms, data aggregation architectures function as such bottlenecks. While SMEs contribute localized data, platforms consolidate cross-actor datasets, enabling algorithmic learning that enhances system-wide performance.

Under these conditions, appropriation is not determined solely by bilateral bargaining. Instead, it is embedded in the design of the aggregation architecture itself. Platform sponsors may design rules governing data access, insight visibility, API permissions, and monetization mechanisms (Chen et al., 2022). These governance mechanisms determine who can access refined analytics, who benefits from cross-actor learning, and how monetization layers (e.g., sponsored placement, data-driven advertising) redistribute surplus.

Thus, AI-centric ecosystems reconfigure appropriation from a negotiation-based process to an architecture-mediated process. Surplus distribution becomes structurally embedded in the platform's control over data infrastructures. This structural embedding lays the groundwork for what this article conceptualizes as data extractivism: the systematic appropriation of value generated from distributed SME data contributions through centralized aggregation and monetization architectures.

2.3 Data Network Effects and Learning Rents

A defining characteristic of AI-driven competition is the compounding nature of data-enabled learning. When machine learning systems improve as data accumulates, early advantages in data scale can translate into persistent performance differentials (Hagiu & Wright, 2020). Unlike traditional network effects, which derive value from user participation alone, data network effects derive value from the informational feedback loops that participation generates.

In multi-actor ecosystems, however, data inputs are not confined to a single firm. SMEs operating on a shared platform collectively generate transactional and behavioral data that feed centralized learning systems. Platforms may pool these heterogeneous data streams to train recommendation models, optimize pricing algorithms, or refine targeting engines. The resulting learning improvements benefit the ecosystem's overall efficiency—but may not proportionally benefit each SME contributor.

This asymmetry introduces the notion of learning rents: returns accruing to actors controlling the aggregation and training architecture rather than to those generating raw data inputs. Because learning improvements compound over time, control over aggregation may yield increasing returns that are difficult for individual SMEs to replicate independently. In such contexts, competitive advantage emerges not from the volume of data generated by a single SME but from the platform's capacity to integrate and leverage cross-SME datasets.

Consequently, SMEs may experience a paradoxical dynamic. Increased participation and data generation can enhance short-term performance through improved matching and visibility. Simultaneously, the same data contributions may strengthen the platform's learning capabilities, thereby reinforcing the platform's relative bargaining position. This dynamic illustrates how data extractivism can coexist with apparent ecosystem growth.

2.4 Platform Governance and Control over Data Flows

Digital platforms are increasingly conceptualized as meta-organizations that orchestrate ecosystems of interdependent actors (Kretschmer et al., 2022). Governance within these systems combines incentive structures and control mechanisms that shape participant behavior (Chen et al., 2022). Importantly, governance extends beyond formal contracts to include algorithmic rules embedded in ranking systems, API architectures, and data access policies.

In AI-centric ecosystems, governance mechanisms determine the terms under which SMEs access data insights, analytics dashboards, and algorithmic optimization tools. Platforms may restrict access to granular performance data, limit portability of historical records, or require paid promotion for enhanced visibility. Such mechanisms can transform data into a quasi-rival resource controlled by the platform sponsor.

Control over data flows thus becomes a central appropriation lever. When platforms design aggregation architectures that centralize learning while limiting complementor-level data access, they shape the distribution of learning rents. Governance opacity may further exacerbate asymmetries by obscuring how data are used and monetized. In the absence of transparent portability and auditability mechanisms, SMEs may lack the information necessary to negotiate alternative arrangements or multi-home effectively.

Therefore, governance design is not merely a coordination device but a structural determinant of rent allocation. Data extractivism, in this view, is not the result of opportunistic behavior but a predictable outcome of centralized architectural control combined with asymmetric data access rights.

2.5 Toward a Theory of Data Extractivism in SME Ecosystems

The integration of RBV, value appropriation theory, data-enabled learning, and platform governance suggests a structural reconfiguration of competitive advantage in AI-centric ecosystems. Competitive rents increasingly derive from control over aggregation architectures rather than from ownership of localized data inputs. Value appropriation becomes architecture-mediated, embedded within governance mechanisms that regulate data access, recombination, and monetization.

For SMEs, this implies that participation in AI-driven platforms entails both opportunity and structural constraint. While digital ecosystems can enhance visibility and efficiency, they may simultaneously stabilize extractive equilibria in which SMEs generate data-driven value without proportionate appropriation. Understanding these dynamics requires moving beyond firm-centric resource models toward ecosystem-aware theories of architectural control.

3. Data Extractivism and Architecture-Mediated Value Appropriation

The preceding discussion establishes that AI-centric ecosystems alter the structural conditions under which competitive advantage and value appropriation emerge. Building on the integration of RBV, value appropriation theory, data-enabled learning, and platform governance, this section develops a mechanism-based conceptual model explaining how data extractivism operates within SME-dominated ecosystems. The model specifies how SME data contributions are transformed through centralized aggregation architectures into learning rents that disproportionately accrue to platform sponsors, and under what boundary conditions this asymmetry stabilizes or attenuates.

At the core of the model lies a fundamental distinction between data generation and data aggregation control. SMEs generate transactional, behavioral, and contextual data as a byproduct of their participation in digital platforms. These data streams are locally embedded and often idiosyncratic. However, the economic value of such data does not arise solely from their existence but from their transformation into predictive intelligence through aggregation, recombination, and machine learning. Control over this transformation process—rather than control over isolated data points—becomes the decisive locus of competitive advantage.

3.1 Aggregation Architecture Control as the Primary Appropriation Mechanism

In AI-centric ecosystems, data aggregation architectures function as bottlenecks in the sense described by ecosystem theory (Jacobides et al., 2018). They are structural nodes where distributed inputs converge and are transformed into system-level capabilities. Platforms design and control these architectures, determining which data are pooled, how they are cleaned and standardized, how models are trained, and how outputs are deployed across the ecosystem.

From an RBV perspective, aggregation architecture control constitutes a higher-order resource: it governs access to and transformation of underlying data inputs. While SMEs may possess valuable data at the transactional level, they typically lack access to cross-SME datasets and the computational infrastructure required for large-scale learning. Consequently, competitive rents increasingly derive from the authority to orchestrate data flows rather than from the possession of isolated data assets.

This structural asymmetry implies that increases in SME data contribution intensity do not automatically translate into proportional increases in SME value capture. Instead, the marginal value of additional data contributions may accrue primarily to the platform if it retains exclusive control over aggregation and model deployment. Therefore:

Proposition 1.
In AI-centric ecosystems, control over the data aggregation architecture predicts value appropriation more strongly than SME-level data generation intensity.

This proposition reframes competitive advantage from a firm-centric to an architecture-centric phenomenon. It suggests that the critical resource is not data per se, but the institutionalized authority to determine how data are pooled, modeled, and monetized.

3.2 Cross-SME Data Pooling and the Emergence of Learning Rents

Data-enabled learning introduces dynamic feedback effects that amplify initial asymmetries. As Hagiu and Wright (2020) demonstrate, performance improvements in AI systems depend on the volume and diversity of training data. In platform ecosystems, data pooling across heterogeneous SMEs increases the richness and generalizability of learning models. Recommendation engines, fraud detection systems, and pricing algorithms improve as they observe broader behavioral patterns across markets and categories.

Importantly, these learning gains are non-rival at the platform level but rival at the SME level. A single SME cannot replicate the cross-sectional diversity available to the platform. Thus, learning improvements derived from cross-SME pooling generate returns that are structurally inaccessible to individual complementors. Over time, these improvements compound, reinforcing the platform's relative informational advantage.

The compounding nature of such learning processes gives rise to what may be termed learning rents: persistent returns accruing to actors who control cumulative learning trajectories. Because SMEs contribute data that feed these trajectories without accessing equivalent cross-actor insights, their relative bargaining position may weaken as the ecosystem matures.

Proposition 2.
Cross-SME data pooling in AI-centric ecosystems generates compounding learning rents that disproportionately accrue to actors controlling the aggregation and model-training architecture.

This mechanism explains how ecosystem growth can coincide with increasing asymmetry in rent distribution. While overall system efficiency may improve, the distribution of surplus may shift toward centralized actors.

3.3 Architecture-Mediated Monetization and Surplus Redistribution

The translation of learning rents into economic surplus depends on monetization design. Platform sponsors frequently embed monetization levers—such as sponsored ranking, data-driven advertising, analytics subscriptions, and dynamic commission structures—within the same aggregation architecture that produces learning advantages. Governance research emphasizes that such design choices combine incentives and control to shape ecosystem behavior (Chen et al., 2022).

As ecosystems evolve from growth-oriented phases to monetization-oriented phases, platforms may intensify extraction mechanisms. Data-driven ranking systems can privilege paid promotion; analytics dashboards may reveal performance gaps without granting access to the underlying data required to close them independently. Because monetization tools are layered atop centralized learning systems, SMEs may find themselves compelled to purchase access to visibility or analytics derived from their own collective data contributions.

This dynamic embeds value appropriation within the platform's architectural design rather than within explicit bargaining processes. Surplus redistribution occurs not through direct price negotiations but through structural control over monetization interfaces.

Proposition 3.
As monetization intensity increases within AI-centric ecosystems, the proportion of surplus derived from SME-generated data that accrues to platform sponsors increases relative to SME-level value capture.

This proposition highlights a phase-dependent dynamic in ecosystem evolution: growth phases may diffuse benefits broadly, whereas monetization phases consolidate returns through architectural levers.

3.4 Lock-In, Data Non-Portability, and the Stabilization of Extractive Equilibria

A critical question concerns the persistence of such asymmetries. Why would SMEs continue participating in ecosystems where value appropriation becomes increasingly skewed? The answer lies in switching costs, data non-portability, and relational lock-in.

The next visual isolates the structural conditions under which asymmetries in AI-centric ecosystems persist over time. Rather than depicting value creation, it focuses on equilibrium stabilization—showing how switching costs, data non-portability, and governance opacity reinforce architecture-mediated appropriation and prevent rebalancing.

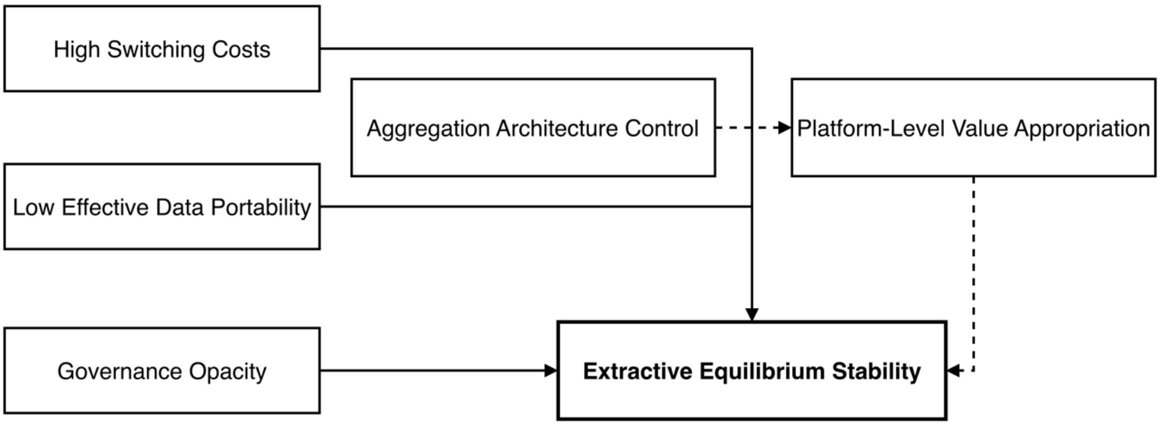

Figure 1. Structural Stabilization of Extractive Equilibria in AI-Centric Ecosystems
Source: Author's conceptualization

Figure 1 reorients the analysis toward persistence rather than initial appropriation. Aggregation architecture control enables platform-level value capture, which feeds into the stabilization of an extractive equilibrium. High switching costs, low effective data portability, and governance opacity reinforce this equilibrium by constraining SME exit and weakening outside options. By visualizing these reinforcing conditions, Figure 1 clarifies why asymmetrical rent distribution can endure even when SMEs remain active ecosystem participants.

Data accumulated within platform infrastructures often lack seamless portability. Historical transaction records, customer reviews, behavioral analytics, and algorithmic performance scores may be difficult to transfer across platforms. Even when formal portability rights exist, technical and procedural barriers can limit effective mobility. High switching costs weaken SMEs' outside options, thereby constraining their bargaining leverage.

From a value appropriation perspective, extractive equilibria emerge when the expected gains from continued participation exceed the costs of exit, even if relative rent shares decline. SMEs may experience declining relative returns yet remain dependent on platform-generated demand and infrastructure. In such settings, centralized control over aggregation architecture interacts with switching costs to stabilize asymmetric surplus distribution.

Proposition 4.
High switching costs and low effective data portability strengthen the stability of extractive equilibria in AI-centric SME ecosystems.

This proposition positions lock-in not merely as a transactional inconvenience but as a structural moderator of rent distribution dynamics.

3.5 SME Capabilities as Moderating Conditions

Although architectural control confers structural advantages, SME heterogeneity remains consequential. SMEs with higher absorptive capacity, stronger data analytics skills, or diversified channel strategies may partially offset extractive pressures. By developing proprietary analytics capabilities, negotiating data-sharing agreements, or multi-homing across platforms, such SMEs can improve their appropriation outcomes.

The following figure isolates the moderating role of SME absorptive capacity within the broader architecture-mediated appropriation process. Rather than depicting structural dominance alone, it specifies how firm-level capability conditions the translation of data contribution into SME-level value capture. This clarifies that ecosystem asymmetry is structured but not fully deterministic.

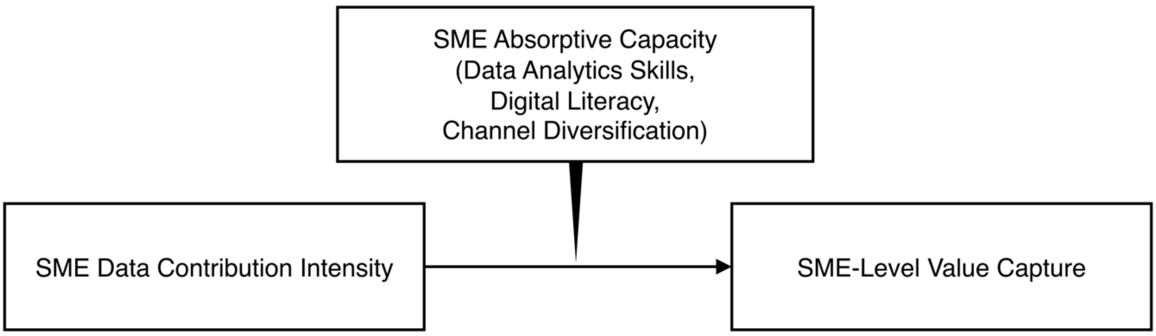

Figure 2. Moderating Role of SME Absorptive Capacity in Architecture-Mediated Ecosystems
Source: Author's conceptualization

Figure 2 specifies the conditional nature of SME value capture within AI-centric ecosystems. While data contribution intensity constitutes a necessary input, its translation into SME-level value capture depends on absorptive capacity—encompassing analytics skills, digital literacy, and channel diversification. By modeling absorptive capacity as a moderator of the primary relationship, Figure 2 clarifies that structural asymmetry does not eliminate firm-level agency but conditions its effectiveness within architecture-mediated competitive environments.

Dynamic capabilities research emphasizes the importance of sensing, seizing, and reconfiguring resources in turbulent environments (Teece, 2007). In AI-centric ecosystems, SMEs capable of interpreting platform-provided analytics critically and integrating external data sources may capture greater value from their participation. Conversely, SMEs with limited digital literacy may become increasingly dependent on platform-generated insights.

Proposition 5.
SME absorptive capacity positively moderates the relationship between data contribution intensity and SME-level value capture in AI-centric ecosystems.

This moderation acknowledges that structural asymmetry does not eliminate agency; rather, it conditions the effectiveness of SME strategic responses.

3.6 Governance Transparency and Appropriation Constraints

Finally, governance transparency influences the extent to which architectural control translates into extractive outcomes. When platforms disclose data usage practices, provide auditability of algorithms, or enable meaningful data portability, asymmetries may be mitigated. Transparent governance reduces informational opacity and may empower SMEs to make informed strategic decisions.

Conversely, opaque governance can obscure the mechanisms through which data contributions are monetized, limiting SMEs' ability to assess the true distribution of surplus. In such cases, architecture-mediated appropriation becomes less visible yet more entrenched.

Proposition 6.
Higher levels of governance transparency weaken the relationship between aggregation architecture control and platform-level value appropriation.

This final proposition introduces an institutional boundary condition, recognizing that regulatory and governance design choices can reshape appropriation dynamics.

3.7 Integrative Logic of the Model

Taken together, the model posits a sequential mechanism: SME data contribution feeds into centralized aggregation architectures; aggregation enables cross-SME learning and the accumulation of learning rents; monetization levers embedded within the same architecture translate learning rents into economic surplus; switching costs and governance opacity stabilize extractive equilibria; and SME capabilities condition firm-level outcomes.

The following visual articulates the core sequential mechanism through which data extractivism operates in AI-centric SME ecosystems. It formalizes the transformation of distributed SME data contributions into architecture-mediated surplus redistribution via centralized aggregation and compounding learning processes.

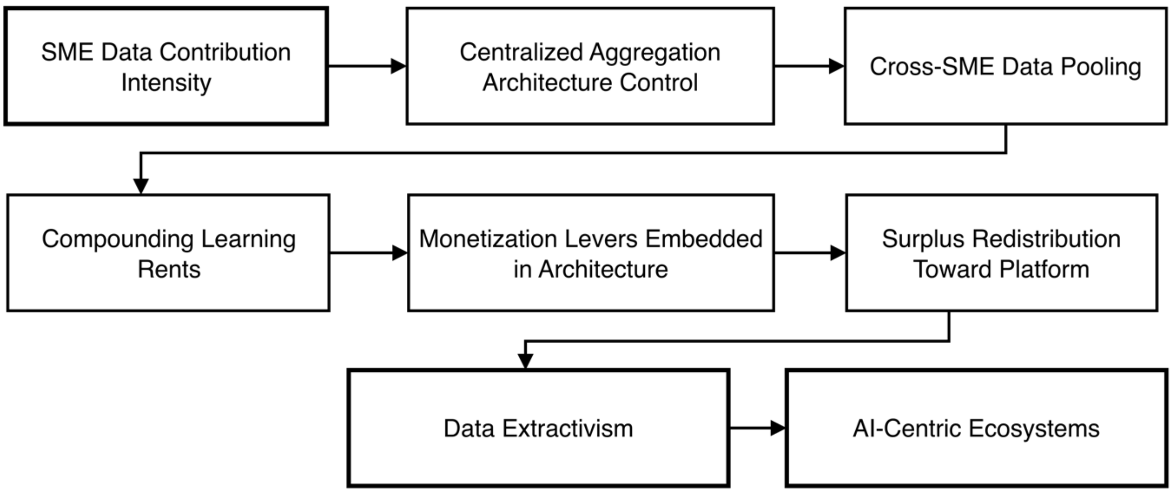

Figure 3. Sequential Mechanism of Architecture-Mediated Value Appropriation
Source: Author's conceptualization

The framework articulated in Figure 3 clarifies the mechanism linking distributed SME data generation to platform-level surplus capture. Data contributions first enter a centralized aggregation architecture, enabling cross-SME pooling and the accumulation of compounding learning rents. These rents are subsequently translated into economic surplus through monetization levers embedded within the same infrastructure, culminating in redistribution toward the platform sponsor. By structuring the model sequentially, Figure 3 makes explicit how value creation and value appropriation become decoupled through architectural control.

The conceptual shift advanced here is from ownership-based competitive advantage to architecture-based competitive advantage. In AI-centric ecosystems, the decisive strategic asset is not the localized data generated by SMEs but the authority to orchestrate, aggregate, and monetize distributed data flows. Data extractivism thus emerges not as a normative accusation but as a structural property of ecosystems characterized by centralized aggregation and asymmetric control rights.

4. Theoretical Repositioning of Competitive Advantage in AI-Centric SME Ecosystems

The conceptual model advanced in this study compels a reconsideration of foundational assumptions in strategic management theory. By foregrounding data extractivism as an architecture-mediated process, the analysis challenges established premises in the resource-based view (RBV), value appropriation theory, and ecosystem strategy. The discussion below situates the framework within these traditions, identifies points of theoretical obsolescence, and articulates how AI-centric ecosystems require conceptual revision.

The table below condenses the article's theoretical repositioning into a structured comparison across the three main streams it engages.

Table 1. Theoretical Repositioning in AI-Centric SME Ecosystems

Theoretical Stream	Classical Assumption	AI-Centric Reconfiguration	Contribution of This Article
Resource-Based View (RBV)	Competitive advantage derives from firm-owned VRIN resources under firm control	Data generation is distributed; aggregation control is centralized	Reframes advantage as architecture-based rather than ownership-based
Value Appropriation Theory	Surplus division determined primarily by bargaining power and outside options	Surplus distribution embedded in platform architecture and governance design	Conceptualizes appropriation as architecture-mediated rather than negotiation-mediated
Ecosystem / Platform Strategy	Platforms coordinate complementors to enable growth and innovation	Cross-SME data pooling generates compounding learning rents concentrated at the architectural center	Introduces learning rents as a structural mechanism of asymmetrical value capture

Source: Developed by the author

Table 1 sharpens the article's theoretical contribution by explicitly contrasting foundational assumptions with their AI-centric reformulation. By presenting the repositioning in structured form, Table 1 makes clear that the contribution lies not in extending each theory incrementally, but in reconfiguring their core premises around aggregation architecture control.

4.1 From Ownership-Based Advantage to Architecture-Based Advantage

Classical RBV posits that competitive advantage arises from firm-specific ownership and control of valuable resources (Barney, 1991; Peteraf, 1993). The logic presumes that rents accrue to those who possess and protect strategic assets through isolating mechanisms. Even dynamic capabilities theory, while extending the RBV to turbulent environments, maintains the firm as the locus of sensing, seizing, and transforming resources (Teece, 2007).

However, AI-centric ecosystems destabilize the alignment between ownership and rent capture. SMEs may generate economically valuable data through transactions and interactions, yet lack authority over how those data are aggregated, recombined, and monetized. In such contexts, resource possession does not guarantee rent appropriation. Control over aggregation architecture becomes the decisive mechanism.

The figure below clarifies the theoretical shift from ownership-based competitive advantage to architecture-based competitive advantage in AI-centric ecosystems. It isolates the locus of rent appropriation under traditional RBV assumptions and contrasts it with the architecture-mediated logic developed in this study. By structurally juxtaposing the two logics, the visual reframes competitive advantage as dependent on aggregation control rather than mere resource possession.

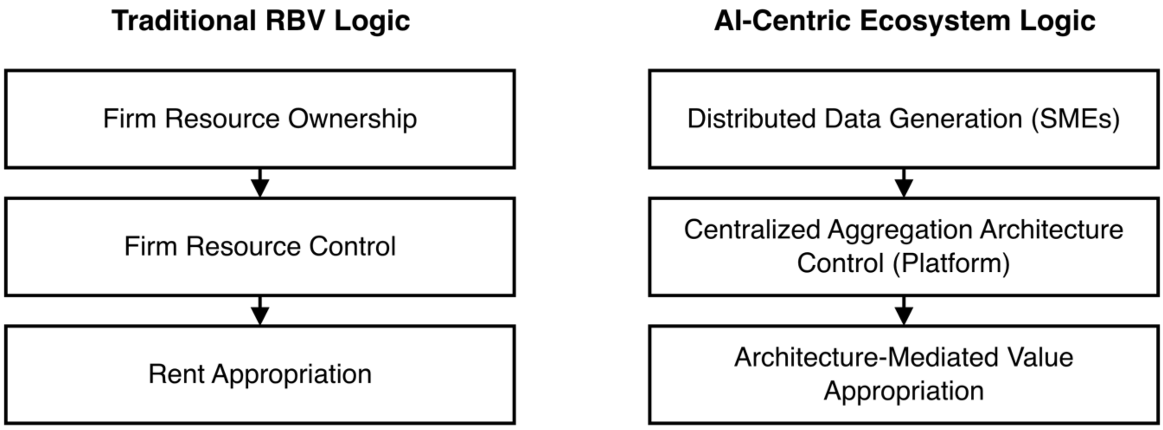

Figure 4. From Ownership-Based to Architecture-Based Competitive Advantage
Source: Author's conceptualization

As illustrated in Figure 4, the left-hand logic reflects the classical RBV assumption that resource ownership and control are co-located within firm boundaries, leading directly to rent appropriation. In contrast, the right-hand structure reorients competitive advantage toward centralized aggregation architecture control within AI-centric ecosystems. Figure 4 therefore establishes the article's core theoretical repositioning: when data generation is distributed but aggregation control is centralized, rent appropriation becomes architecture-mediated rather than ownership-based.

This shift exposes a latent limitation in RBV: its implicit assumption of boundary congruence between resource generation and resource control. As Jacobides, Cennamo, and Gawer (2018) argue, ecosystems fragment production and reallocate control over critical bottlenecks. In AI-driven settings, data aggregation infrastructures function as such bottlenecks, concentrating strategic authority in platform sponsors.

Recent work on data-enabled learning reinforces this reorientation. Hagiu and Wright (2020) demonstrate that AI learning advantages intensify with cumulative data. Yet their models implicitly assume firm-level integration of data flows. In multi-actor ecosystems, centralized platforms—not individual SMEs—accumulate cross-sectional datasets and orchestrate model training. Thus, competitive advantage migrates from resource ownership to architectural orchestration.

The implication is profound: VRIN criteria must be reconsidered. In AI ecosystems, rarity and inimitability derive less from exclusive possession of data and more from exclusive control over aggregation pathways. The relevant strategic resource becomes the governance architecture itself.

4.2 Rethinking Value Appropriation: From Bargaining to Structural Mediation

Value appropriation theory traditionally emphasizes bargaining power, outside options, and contractual arrangements as determinants of surplus division (Brandenburger & Stuart, 1996). In vertically structured markets, rent distribution reflects negotiated outcomes among identifiable actors.

AI-centric ecosystems introduce a structural mediation layer that weakens the explanatory sufficiency of bargaining-based models. Surplus distribution is increasingly embedded within algorithmic infrastructures—ranking systems, data dashboards, monetization APIs—rather than negotiated explicitly. Governance research shows that platform sponsors shape complementor behavior through incentive and control mechanisms embedded in design (Chen et al., 2022). These mechanisms influence access to visibility, analytics, and monetization opportunities.

The critical insight is that architecture shapes outside options ex ante. When switching costs are high and data portability limited, SMEs' bargaining leverage is structurally constrained before negotiation occurs. OECD analyses of digital markets highlight how limited portability and interoperability reinforce concentration (OECD, 2024). Thus, value appropriation becomes architecture-mediated rather than bargaining-determined.

This challenges a core premise of value-based strategy: that surplus division primarily reflects actor-level negotiation. In AI ecosystems, surplus division is coded into the design of aggregation infrastructures. Data extractivism, therefore, should be understood not as opportunistic exploitation but as a structural outcome of centralized architectural control.

4.3 Ecosystem Growth and the Paradox of Learning Rents

Platform strategy literature frequently emphasizes growth, innovation, and complementor participation as mutually reinforcing dynamics (Kretschmer et al., 2022). Digital ecosystems are often portrayed as enabling SMEs by reducing entry barriers and expanding market reach.

While such narratives capture early-stage growth dynamics, they understate the distributional consequences of cross-actor learning. As data pooling expands, AI systems improve system-wide efficiency. However, the benefits of learning are not evenly distributed. Cross-SME data aggregation generates compounding learning rents that are difficult for individual SMEs to replicate.

This dynamic introduces a paradox: ecosystem expansion may simultaneously enhance aggregate welfare and intensify asymmetry in rent capture. SMEs may experience improved short-term performance due to better matching and targeting, yet their relative share of ecosystem surplus may decline over time. Learning rents accumulate at the architectural center.

Such dynamics resonate with recent debates on platform power and attention rents, which argue that centralized infrastructures extract value from distributed participation (Gawer, 2021). However, prior work has not systematically integrated learning effects into value appropriation theory. The present framework extends ecosystem scholarship by theorizing how AI-based learning trajectories compound structural asymmetries.

4.4 Revisiting SME Strategy Under Extractive Architectures

Traditional SME strategy research emphasizes agility, niche specialization, and relational embeddedness as sources of advantage. Digital platforms are often framed as leveling mechanisms that empower small firms by providing scalable infrastructure.

The present analysis complicates this view. While platforms provide access to markets and analytics, they may simultaneously centralize learning benefits and monetization control.

SMEs with limited absorptive capacity may become dependent on platform-generated insights without accessing the underlying data necessary for independent learning.

Dynamic capabilities theory suggests that firms can reconfigure resources to adapt (Teece, 2007). Yet adaptation presumes access to relevant resources. In AI-centric ecosystems, the resource that matters most—aggregation architecture control—lies outside SME boundaries. Thus, SME strategic responses are conditioned by structural constraints that earlier theory did not anticipate.

This does not imply SME passivity. Multi-homing, collective bargaining, and proprietary data strategies may mitigate extractive pressures. However, such strategies require capabilities and institutional conditions that are unevenly distributed. Governance transparency and regulatory frameworks further shape these possibilities.

4.5 Critical Implications for Contemporary Strategy Theory

Taken together, the findings suggest that strategy theory must move beyond firm-centric models in digitally mediated environments. Competitive advantage increasingly derives from positional authority within data infrastructures rather than from isolated firm resources. Appropriation is structurally mediated by architectural design. Learning rents compound asymmetry over time.

These insights challenge the continued reliance on static VRIN criteria, bilateral bargaining logic, and growth-centric ecosystem narratives. AI-centric markets require a re-theorization of advantage as architecture-dependent, appropriation as structurally embedded, and SME performance as ecosystem-conditioned.

In this sense, data extractivism is not merely a descriptive phenomenon but a theoretical lens for understanding how digital infrastructures redistribute economic rents. By integrating RBV, value appropriation theory, and platform governance, this study offers a reframed account of competitive advantage suited to the realities of AI-driven ecosystems.

5. Boundary Conditions, Future Research, and Implications

The theoretical repositioning advanced in this article is intentionally structural. However, the dynamics of data extractivism and architecture-mediated appropriation are not universal across all digital contexts. Their intensity and consequences depend on identifiable boundary conditions and institutional configurations. Furthermore, the conceptual reframing offered here opens several avenues for empirical testing and theoretical refinement. Finally, the implications for SME strategy and public policy require careful articulation, particularly given the centrality of digital platforms in contemporary economic systems.

5.1 Boundary Conditions of Data Extractivism in SME Ecosystems

The proposed framework is most applicable in ecosystems characterized by high centralization of data aggregation and strong AI integration in core decision processes. In markets where algorithmic matching, ranking, and personalization are peripheral rather than central, aggregation architecture control may exert weaker effects on rent distribution. Thus, the degree of AI centrality constitutes a primary boundary condition.

Industry data intensity further moderates extractive dynamics. In sectors where transactional and behavioral data are highly predictive—such as e-commerce, mobility platforms, or digital services—cross-actor pooling generates stronger learning rents. In contrast, industries with lower data richness or weaker feedback loops may exhibit attenuated asymmetries. The compounding logic of learning rents depends on the strength of informational returns to scale (Hagiu & Wright, 2020).

Market maturity also shapes appropriation patterns. During early growth phases, platform sponsors may subsidize complementors to stimulate participation. As ecosystems mature, monetization pressures intensify, and appropriation levers become more salient. Thus, extractive equilibria are more likely to stabilize in later lifecycle stages, when network effects have solidified and switching costs have increased (Kretschmer et al., 2022).

Regulatory environments introduce additional constraints. Jurisdictions with enforceable data portability, interoperability mandates, or transparency requirements may weaken architecture-mediated appropriation. However, as recent analyses suggest, portability rights alone do not guarantee effective mobility if technical and procedural barriers persist (OECD, 2024). Therefore, institutional strength and enforcement capacity moderate the persistence of extractive equilibria.

Finally, SME heterogeneity constitutes a firm-level boundary condition. SMEs with strong absorptive capacity, diversified channel strategies, or collective bargaining mechanisms may partially mitigate extractive pressures. Conversely, digitally constrained SMEs operating in single-homing environments are more likely to experience architecture-dependent disadvantage.

5.2 Future Research Directions

The conceptual model proposed here invites systematic empirical examination. First, quantitative studies could assess whether aggregation architecture control predicts value appropriation more strongly than firm-level data intensity. Multi-level panel datasets linking SME performance to platform-level governance changes would enable testing of Propositions 1–3.

Second, longitudinal research could examine lifecycle transitions from growth-oriented to monetization-oriented ecosystem phases. Such studies would clarify whether surplus distribution shifts systematically as AI integration deepens and learning rents accumulate.

Third, comparative cross-country analyses could explore how regulatory frameworks moderate extractive equilibria. Differences in data portability enforcement, algorithmic transparency mandates, or antitrust interventions provide natural variation for testing Proposition 6.

Fourth, microfoundational research should investigate SME strategic responses. Experimental or survey-based studies could examine how absorptive capacity, digital literacy, and multi-homing strategies influence SME value capture within centralized architectures.

The table summarizes the six propositions into a structured analytical overview. It clarifies the core mechanisms, predicted outcomes, and relevant boundary conditions without repeating the full argumentation. In essence, the table helps readers understand the internal logic of the framework in a concise and systematically organized manner.

Table 2. Propositional Architecture of Data Extractivism in AI-Centric Ecosystems

Proposition	Core Mechanism	Predicted Outcome	Key Boundary Conditions
P1	Aggregation architecture control outweighs SME-level data generation intensity	Platform-level value appropriation increases relative to SME capture	Degree of AI centrality; data intensity of industry
P2	Cross-SME data pooling enables compounding learning trajectories	Learning rents accrue disproportionately to aggregation controllers	Strength of informational returns to scale
P3	Monetization levers embedded in aggregation	Surplus redistribution toward platform sponsor	Ecosystem lifecycle stage; governance design

	architecture translate learning rents into surplus	intensifies in monetization phase	
P4	High switching costs and low effective data portability constrain SME exit	Extractive equilibria stabilize over time	Market maturity; interoperability constraints
P5	SME absorptive capacity conditions translation of data contribution into firm-level returns	Higher SME value capture under equivalent structural conditions	Digital literacy; analytics capability; channel diversification
P6	Governance transparency limits opacity in data use and monetization	Weaker linkage between aggregation control and platform-level appropriation	Regulatory enforcement strength; auditability mechanisms

Source: Developed by the author

Table 2 provides a compact representation of the article's causal architecture. By aligning each proposition with its mechanism, outcome, and boundary conditions, Table 2 enhances conceptual precision and demonstrates that the framework is mechanism-driven rather than merely descriptive.

Finally, future theoretical work may integrate institutional theory and political economy perspectives more explicitly. While this article focuses on strategic implications, broader societal consequences of data extractivism warrant deeper interdisciplinary analysis.

5.3 Managerial Implications for SMEs

The framework developed here suggests that SMEs operating within AI-centric ecosystems must reconsider the basis of competitive advantage. Rather than assuming that data generation alone strengthens bargaining position, SMEs should assess who controls aggregation architectures and how learning rents are distributed.

Strategically, SMEs may benefit from investing in proprietary data analytics capabilities that reduce reliance on platform-provided insights. Multi-homing across platforms can mitigate lock-in risks and preserve outside options, although feasibility varies by industry. Collective coordination among SMEs—such as shared data cooperatives or negotiated data-sharing agreements—may also rebalance appropriation dynamics.

Moreover, SMEs should treat platform governance transparency as a strategic variable. Understanding how ranking, monetization, and data policies operate is essential for evaluating long-term rent capture rather than short-term performance gains.

5.4 Policy Implications for Digital Market Governance

From a policy perspective, the findings suggest that competition concerns in AI-centric markets cannot be assessed solely through price effects or market share concentration. Architectural control over data aggregation and learning infrastructures may constitute a structural source of market power even when participation appears open.

Effective data portability, interoperability mandates, and transparency requirements can reduce asymmetry in aggregation control. However, regulatory design must address practical enforceability, not merely formal rights. Policies that enhance SME access to aggregated analytics or mandate clearer disclosure of data usage may mitigate extractive equilibria.

Importantly, the goal of policy intervention should not be to undermine ecosystem efficiency but to ensure that learning rents do not systematically concentrate in ways that stifle SME viability and long-term innovation diversity.

6. Conclusion

This study has advanced a structural reinterpretation of competitive advantage in AI-centric SME ecosystems. By integrating the resource-based view, value appropriation theory, data-enabled learning, and platform governance research, the analysis demonstrates that in digitally mediated markets, competitive rents increasingly derive not from resource ownership but from control over data aggregation architectures. The central argument is that AI-driven ecosystems decouple data generation from data control, thereby destabilizing the classical alignment between resource possession and rent appropriation.

In traditional strategy theory, firms capture value from assets they own, protect, and deploy. However, when SMEs operate within centralized AI infrastructures, the decisive strategic asset becomes the authority to orchestrate, aggregate, and monetize distributed data flows. Aggregation architectures function as ecosystem bottlenecks, enabling cross-SME data pooling that produces compounding learning rents. These rents accrue disproportionately to actors controlling the training and deployment infrastructure. Value appropriation thus becomes architecture-mediated rather than negotiation-based.

The concept of data extractivism proposed here should not be interpreted normatively but analytically. It describes a structural condition in which distributed actors contribute data that enhance system-wide learning while centralized actors capture a disproportionate share of the resulting surplus. Such dynamics may coexist with ecosystem growth and improved efficiency, generating a paradox in which aggregate value creation increases even as relative SME value capture declines.

By repositioning competitive advantage as architecture-dependent, this article extends RBV beyond its ownership-centric assumptions and reframes value appropriation as structurally embedded in governance design. It also contributes to ecosystem strategy by highlighting how AI-driven learning trajectories amplify asymmetries over time. For SMEs, the findings imply that participation in digital platforms entails strategic trade-offs between short-term market access and long-term appropriation constraints.

More broadly, the study underscores the necessity of ecosystem-aware strategy theory in an era where data flows transcend firm boundaries and AI infrastructures shape market outcomes. As digital platforms become increasingly central to economic coordination, understanding who controls aggregation architectures—and how that control structures rent distribution—will remain critical for both theory and practice.

In AI-centric markets, the locus of advantage is shifting. Strategy must therefore move from asking who owns valuable resources to asking who governs the architectures that transform distributed data into scalable intelligence.

References

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Brandenburger, A. M., & Stuart, H. W. (1996). Value-based business strategy. *Journal of Economics & Management Strategy, 5*(1), 5–24. https://doi.org/10.1111/j.1430-9134.1996.00005.x

Chen, Y., Pereira, I., & Patel, P. C. (2022). Decentralized governance of digital platforms. *Organization Science, 33*(6), 2223–2244. https://doi.org/10.1287/orsc.2022.1594

Gawer, A. (2021). Digital platforms and ecosystems: Remarks on the dominant organizational forms of the digital age. *Innovation: Organization & Management, 23*(1), 110–124. https://doi.org/10.1080/14479338.2020.1857362

Hagiu, A., & Wright, J. (2020). When data creates competitive advantage. *Harvard Business Review, 98*(1), 94–101.

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal, 39*(8), 2255–2276. https://doi.org/10.1002/smj.2904

Kretschmer, T., Leiponen, A., Schilling, M., & Vasudeva, G. (2022). Platform ecosystems as meta-organizations: Implications for platform strategies. *Strategic Management Journal, 43*(3), 405–424. https://doi.org/10.1002/smj.3250

OECD. (2024). *Data portability, interoperability and competition in digital markets*. OECD Publishing.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal, 14*(3), 179–191. https://doi.org/10.1002/smj.4250140303

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210. https://doi.org/10.5465/amr.2018.0072

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance. *Strategic Management Journal, 28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Verhoef, P. C., Broekhuizen, T., Bart, Y., Bhattacharya, A., Dong, J. Q., Fabian, N., & Haenlein, M. (2021). Digital transformation: A multidisciplinary reflection and research agenda. *Journal of Business Research, 122*, 889–901. https://doi.org/10.1016/j.jbusres.2019.09.022