Economics Bulletin

Volume 46, Issue 1

The sectoral concentration of data capitalism: Evidence from Latvia

Ritvars Mētra
Latvia University of Life Sciences and Technologies

Abstract

This paper examines whether “data-capitalist” firms are concentrated in specific sectors within a small, open economy. A novel classification tool, the Data-Capitalism Maturity Framework (DCMF), is used to measure the intensity of data capitalism across 101 of Latvia's most valuable firms, which are mapped into four macro-sectors. The association between sector and data-capitalism tier is tested using exact methods with Holm adjustments, while robustness to firm size is assessed with the Cochran–Mantel–Haenszel test. Within each sector, performance metrics (turnover, EBITDA) are compared across tiers using Kruskal–Wallis and Holm-adjusted Mann–Whitney tests. Short-term competitive mobility is evaluated via changes in corporate rankings (ΔRank, 2023–2024). A strong concentration of Primary-tier firms is found in the IT & Telecommunications sector, a result that holds after conditioning on firm size. In contrast, differences in the Financial Services sector do not remain significant after statistical adjustment. Across all sectors, no systematic performance premium for higher-tier firms is observed. Competitive rank mobility does not significantly differ by tier. The study concludes that data capitalism in Latvia is highly sector-specific but is not yet associated with broad, measurable firm-level performance gains. These findings suggest that pathways into the data economy are sector-dependent, challenging the narrative of a uniform digital transformation.

The author report there are no competing interests to declare. The author used ChatGPT 4.0 by OpenAI to refine the language and clarity of this work. Afterward, the author reviewed and edited the content as needed and take full responsibility for the content of the final publication. No funding was received.
Citation: Ritvars Mētra, (2026) "The sectoral concentration of data capitalism: Evidence from Latvia", *Economics Bulletin*, Volume 46, Issue 1, pages 219-231
Contact: Ritvars Mētra - sz19002@lbtu.lv.
Submitted: October 09, 2025. **Published:** March 30, 2026.

1. Introduction

The increasing use of data as a core economic resource is reshaping modern economies, giving rise to a phenomenon known as Data Capitalism (DC). This can be understood as a socio-economic paradigm where the extraction, commodification, and exploitation of data become an important engine of value creation and profit (Bolin, 2022). While the transformative potential of DC is clear, much of the existing discourse focuses on large economies or global tech giants. This overlooks the nuanced ways different industries within smaller nations are adapting to data-centric models. This paper addresses this gap by asking: Are there specific sectors or industries in Latvia that have been disproportionately affected by the rise of data capitalism?

To answer this question, I develop and apply a novel classification tool, the Data-Capitalism Maturity Framework (DCMF), to measure the intensity of DC across 101 of Latvia's most valuable firms during 2023-2024. Firms are categorized into four tiers—from "Primary" data capitalists, where data is the core asset, to "Non-Data Capitalists."

The results show a strong and sector-specific pattern: Primary data capitalism in Latvia is overwhelmingly concentrated in the IT & Telecommunications sector. This concentration is not an artifact of firm size and persists after controlling for it. In contrast, sectors like Manufacturing & Consumer Trade and Financial Services exhibit more diversified and less distinct profiles. I find no evidence that a higher DC tier consistently translates into superior short-term financial performance or competitive mobility in our sample. These findings suggest that pathways into the data economy are highly sector-dependent and challenge the narrative of a uniform digital transformation.

2. Theory development

The shift toward a data-centric economy is heterogeneous across industries. Sectoral structures, competition, and institutions shape both the incentives and feasible trajectories of adoption (Bresnahan et al., 2002). A firm's ability to extract, process, and monetize data depends on its resource base and dynamic capabilities (Teece, 2018). Therefore systematic cross-sector variation in the prevalence and intensity of Data Capitalism (DC) within Latvia is expected. Data-driven strategies are no longer confined to consumer tech: business-to-business models increasingly hinge on data for value creation and capture (Ritala et al., 2024). In manufacturing, "servitization" (smart products, outcome-based services) enables continuous usage data collection and analytics, improving design, reliability, and customer experience (Gebauer et al., 2020; Kowalkowski et al., 2015; Mosch et al., 2022). Such shifts typically deepen data dependency without fully displacing the physical product as the core value proposition—consistent with Tertiary or Secondary DC tiers rather than Primary.

Building on this logic, I derive sector-specific, performance, mobility, and size-robustness hypotheses.

2.1 Sectoral concentration of data capitalism

IT & Telecommunications are "digitally native": products (software, platforms, connectivity) and business models are intrinsically data-based, often scaled through network effects and direct data monetization (Alshawawreh et al., 2024; Banda et al., 2022; Parker et al., 2017; Ritala et al., 2024). Data monetization is strategic, not ancillary, making this sector a natural locus for Primary DC. I propose that IT & Telecommunications have a higher share of Primary data capitalists than other industries *(H1a)*.

Manufacturing & Consumer Trade remain centered on physical goods and distribution. Industry 4.0, smart retail, and predictive maintenance intensify data use but typically as support to the product core (Mostaghel et al., 2022; Porter & Heppelmann, 2014; Shankar et al., 2021). This implies fewer Primary models and a tilt toward other tiers. Therefore I assume that Manufacturing & Consumer Trade are less likely to be Primary than IT & Telecom, and more concentrated in non-Primary tiers (*H1b*).

Financial Services are information-intensive and long reliant on data (risk, scoring, personalization), with FinTech accelerating analytics adoption; yet heavy regulation and legacy systems can dampen radical data-first models (Arner et al., 2015; Babina et al., 2025; Chen et al., 2012; Colangelo, 2024). I expect a diversified tier profile with a large Tertiary share but no outsized Primary presence (H1c).

2.2 Data capitalism and financial performance

From the resource-based view, valuable, rare, and hard-to-imitate assets underpin superior performance (Barney, 1991). Data—paired with analytic and execution capabilities—meets these criteria (McAfee & Brynjolfsson, 2012). By definition, higher DC tiers embed data more deeply in strategy and operations, potentially improving efficiency and market responsiveness. Still, the transition to a data centred business models is not cost-neutral and rarely yields immediate returns. The economics literature has long identified a "productivity paradox," where massive investments in information technology do not immediately translate into measurable productivity or profit gains (Brynjolfsson, 1993). This phenomenon is often explained by the "Productivity J-Curve" (Brynjolfsson et al., 2018). According to this theory, firms at the forefront of technological shifts must invest heavily in "intangible capital"—including proprietary data pipelines, organizational restructuring, and highly specialized human capital. For a firm in the "Primary" or "Secondary" DC tier, these necessary investments are often recorded as current expenses (e.g., R&D, specialized labor, data infrastructure costs), which can suppress current EBITDA and turnover growth. This creates a "performance trough" where the firm is building significant future "option value" that is not yet reflected in its short-term financial statements. In contrast, "Tertiary" or "Non-Data" capitalists may show higher current margins simply because they are not bearing the heavy "transformation costs" associated with the data economy. Therefore, while I expect that within sector groups, financial performance differs across DC_Tier (H2). The relationship may be non-linear due to these masked investment costs.

2.3 Data capitalism and competitive mobility

Dynamic capabilities—sensing, seizing, transforming—enable firms to reconfigure in changing environments (Teece, 2018). Higher-tier data capitalists should exhibit stronger data-driven agility and thus greater upward movement relative to peers. I operationalize mobility as $\Delta\text{Rank} \equiv TOP2023 - TOP2024$ (positive = improvement), and assume that Higher DC_Tier categories show greater upward mobility (ΔRank) (H3).

2.4 The role of firm size

Large firms possess capital and talent to invest in data infrastructure and advanced analytics (McElheran et al., 2024; Oliveira et al., 2014). Apparent Primary **c**oncentration could therefore be a size artifact. To establish DC as a distinct strategic phenomenon, I test independence from size and stratified robustness and propose that DC_Tier is independent of company size (turnover quartiles) (*H4a) and* The IT & Telecom Primary overrepresentation holds after stratifying by size *H4b.*

3. Methodology and Data

3.1 Data

The empirical dataset comprises a cross-section of 101 Latvian firms observed in 2023–2024, sourced from the "TOP 101 of Latvia's Most Valuable Enterprises" list published by Leading Latvian corporate finance company “Prudentia” in collaboration with official stock exchange Nasdaq Riga (top101.lv). This list provides a robust sample of the most significant firms in the national economy. Financial and competitive ranking data were also sourced from this dataset.

3.2 The Data-Capitalism Maturity Framework (DCMF)

Data capitalism can be seen as a scale where non data capitalism is at the one end and primary data capitalists at the another. For this classification firms’ business model should be analyzed in four dimensions - Data Extraction, Data Commodification, Data Exploitation, Data Centrality. The Data-Capitalism Maturity Framework (DCMF) serves as the primary analytical tool for this study. A crucial distinction in the DCMF is between data as a support tool and data as a core asset. While 'visibility bias' may mask internal data-driven efficiencies in sectors like Manufacturing (e.g., supply chain optimization), such activities—by definition—align with the Secondary or Tertiary tiers. For a firm to be classified as 'Primary,' data must be the central engine of value creation and the primary output or commodified resource. Since Primary data capitalism typically involves external-facing products (APIs, platforms, data-monetization), it is inherently more visible in corporate disclosures, anunced job roles, product descriptions, and annual reports than internal-only process optimizations.

The DCMF quantifies the intensity and centrality of data-driven activities across four weighted dimensions: Data Extraction (15%), Data Commodification (35%), Data Exploitation (25%), and Value Creation Through Data (25%). Each dimension is scored on a four-point ordinal scale (0–3), ranging from Minimal/None to Advanced/Extensive, based exclusively on verifiable, publicly available information, including corporate websites, annual reports (from the past three years), product descriptions, official press releases, reputable industry analyses, and relevant job postings. To ensure methodological rigor, the framework mandates conservative scoring, assigning a score of zero in the absence of clear evidence, thereby mitigating speculative assessments.

The classification process follows a sequential algorithm to ensure consistency and reproducibility. First, a score of zero in the Data Extraction dimension automatically classifies an entity as Tier 0 (Non-Data Capitalist), halting further evaluation. For entities with non-zero Data Extraction scores, a composite DCMF Index is calculated as follows:

$$\text{DCMF\_Index\_Raw} = (\text{Score}_{\text{Ext}} \times 0.15) + (\text{Score}_{\text{Com}} \times 0.35) + (\text{Score}_{\text{Exp}} \times 0.25) + (\text{Score}_{\text{Val}} \times 0.25)$$

$$\text{DCMF\_Index\_Score} = \text{Round}\left(\frac{\text{DCMF\_Index\_Raw}}{3.0} \times 100\right)$$

Entities achieving scores of ≥2 in both Data Extraction and Data Commodification are classified as Tier 3 (Primary Data Capitalist), reflecting advanced data-driven business models. For remaining entities, a DCMF Index Score ≥51 results in classification as Tier 2 (Secondary Data Capitalist), while a score ≤50 designates Tier 1 (Tertiary Data Capitalist). Confidence scores (1–3, Low to High) accompany each dimension’s assessment to reflect the robustness of the evidentiary base, enhancing transparency.

The classification process involved a mixed-methods approach integrating qualitative analysis with a "Human-in-the-Loop" (HITL) - AI-assisted methodology. HITL refers to a system or procedure in which a person is actively involved in the operation, monitoring or decision-making of an automated system (Stryker, 2025). First, I conducted a systematic review of public documents—corporate websites, annual reports, job postings (2021–2023), and press releases for each firm. To mitigate the risk of 'signaling bias'—where firms market 'data-centricity' without substance—the DCMF triangulation process included an analysis of job postings. Unlike marketing materials, recruitment data serves as a proxy for actual resource allocation. For each firm "intelligence folders" with annual reports and relevant public documents was created.

Then I manually analyzed and categorized firms into DC tiers using the DCMF protocol. Following preliminary categorization, a validation step was introduced using two large language models (LLMs): Google's Gemini 2.5 Pro and OpenAI's ChatGPT-o3. Each model was provided with the collected public data for each firm and tasked with performing the same classification using a standardized prompt (see Annex A). In instances where the author's initial classification differed from one or both of the LLM-generated classifications, a reconciliation process was initiated. This involved gathering supplementary public information and a thorough re-analysis of the firm against the DCMF criteria. The final classification decision in all cases rested with the human researcher, ensuring that the AI models served as a robust tool for systematic review and validation rather than as autonomous classifiers.

The DCMF provides a systematic, transparent, and replicable methodology for assessing data capitalism maturity, facilitating comparative analysis across Latvia's corporate landscape. However, its reliance on publicly available information may underestimate proprietary or undisclosed data practices, potentially skewing classifications. Additionally, the framework's static assessment approach may not fully capture the dynamic evolution of corporate data strategies. Despite these limitations, the DCMF offers a robust foundation for analyzing the influence of data capitalism on Latvia's regional business landscape, enabling evidence-based insights into its socio-economic implications.

3.3 Analysis

Non-parametric statistics to assess the relationship between sectoral classification and DC tier was used. The primary analysis relies on Pearson's chi-square tests for association between sector and DC tier. To ensure the key finding for the IT & Telecom sector is not an artifact of firm size, I employ a Cochran–Mantel–Haenszel test, stratifying by turnover quartiles.

4. Results

The classification of the 101 firms yielded 10 Primary, 8 Secondary, 43 Tertiary, and 40 Non-Data Capitalist (NDC) firms. Table 1 shows the distribution of these tiers across the major economic sectors.

Table I: DC-Tier Composition by Sector Group (N=101)

DC_Tier	IT & Telecom	Manufacturing & Consumer Trade	Financial Services	Other Industries	Grand Total
Primary	6	1	2	1	10
Secondary	0	1	2	5	8
Tertiary	2	19	6	16	43
NDC	1	20	1	18	40
Sector Total	9	41	11	40	101

A clear pattern stands out. The IT & Telecommunications sector is the locus of Primary data capitalism in Latvia, accounting for 60% of all Primary firms despite representing less than 9% of the firms in the sample. In contrast, the large Manufacturing & Consumer Trade sector has only one Primary firm and is almost evenly split between Tertiary and NDC tiers. To address the heterogeneity within the broad 'Manufacturing & Consumer Trade' sector (N=41), a descriptive breakdown of the sub-sectors was performed. This reveal distinct trajectories of data adoption between industrial producers and retailers. As seen in table 2, in the Tertiary Tier (19 firms), a clear majority (12 firms, or 63%) are engaged in Consumer Trade.

Table II: Manufacturing & Consumer Sub - Sector Breakdown (N=41)

Industry	Primary	Secondary	Tertiary	NDC	Total
Consumer Trade / Retail	0	0	12	6	18
Manufacturing	1	1	6	14	22
Construction / Other Industrial	0	0	1	0	1
Total	1	1	19	20	41

Internal polarization within Manufacturing & Consumer Trade highlights an important asymmetry in the pathways through which data capitalism diffuses across traditional sectors. Retail-oriented firms operate at the interface between producers and final consumers, where data are generated endogenously through transactions, loyalty programs, and digital channels, making data extraction a relatively low-cost complement to existing operations. This structural proximity to consumer-generated data consequently lowers the strategic and organizational barriers to redesigning their business models around data-driven value creation, thereby accelerating the integration of data as a central operational and competitive asset.

In contrast, Conversely, the Non-Data Capitalist (NDC) tier for this sector (20 firms) is dominated by Heavy Manufacturing and Industrial Processing (14 firms, or 70%). These are, for the most part, traditional export-oriented industries such as timber, glass fiber, and food processing. For these entities, the core value proposition remains strictly centered on physical production and material transformation. Heavy manufacturing and industrial processing firms are embedded in capital-intensive production systems where value creation is dominated by scale, physical inputs, and process efficiency, and where data—while potentially valuable for internal optimization—rarely constitute a tradable or monetizable asset.

This structural divergence suggests that sector-level classifications may obscure distinct data-capitalist trajectories within broad industrial categories: consumer-facing segments may advance toward data-enabled hybrid models, while upstream manufacturing cores remain anchored in industrial capitalism. Consequently, the modest aggregate deviation observed for Manufacturing & Consumer Trade should be interpreted not as an incipient sector-wide transition, but as the outcome of compositional shifts driven by retail sub-sectors rather than by a transformation of manufacturing itself.

The Financial Services sector shows a diversified profile, with a majority in the Tertiary tier but no disproportionate concentration in any single category compared to the rest of the economy.

The specific claim that IT & Telecom are disproportionately represented as Primary (H1a) was then examined using a 2×2 exact test of Primary vs non-Primary between IT & Telecom and all other sectors. A very large overrepresentation was detected (Fisher's exact $p \approx 0.000093$, Holm-adjusted $p \approx 0.00037$; $OR \approx 36.75$, 95% CI [5.51, 245.05]), with a complementary underrepresentation in Tertiary (Holm-adjusted $p \approx 0.0185$; $OR \approx 0.10$). Full per-tier contrasts are provided in table below.

Table III: Per-tier Fisher exact tests with Holm–Bonferroni (within sector)].

Sector group (vs Others)	DC_Tier	a (Group & Tier)	b (Group & !Tier)	c (Others & Tier)	d (Others & !Tier)	Odds Ratio	OR 95% CI Low	OR 95% CI High	Fisher p (raw)	Fisher p (Holm-adjusted within group)
IT & Telecom	Primary	6	3	4	88	44	7.956339	243.328	1.25E-05	4.98E-05
IT & Telecom	Secondary	0	9	8	84	0.52322	0.027938	9.798826	1	1
IT & Telecom	Tertiary	2	7	41	51	0.355401	0.070026	1.803762	0.294463	0.588926
IT & Telecom	NDC	1	8	39	53	0.169872	0.020398	1.414686	0.083425	0.250274
Manufact. & Consumer Trade	Primary	1	40	9	51	0.141667	0.017226	1.1651	0.045427	0.181707
Manufact. & Consumer Trade	Secondary	1	40	7	53	0.189286	0.022378	1.60106	0.13754	0.41262
Manufact. & Consumer Trade	Tertiary	19	22	24	36	1.295455	0.580784	2.889546	0.545649	0.545649
Manufact. & Consumer Trade	NDC	20	21	20	40	1.904762	0.843688	4.30031	0.14843	0.29686
Financial Services	Primary	2	9	8	82	2.277778	0.417993	12.41235	0.297884	0.595769
Financial Services	Secondary	2	9	6	84	3.111111	0.545116	17.75587	0.209776	0.629328
Financial Services	Tertiary	6	5	37	53	1.718919	0.488102	6.053414	0.521407	0.521407
Financial Services	NDC	1	10	39	51	0.130769	0.016053	1.065238	0.046227	0.18491

The proposition that Manufacturing & Consumer Trade are predominantly Tertiary (H1b) was not supported after multiplicity control. Although the sector's overall association with DC_Tier in the 4×2 comparison was modest but statistically detectable—$\chi^2(3) = 8.20$, $p = 0.042$, Cramér's $V = 0.285$—the targeted Primary vs non-Primary contrast did not remain significant once adjusted (OR = 0.142, 95% CI [0.017, 1.165]; Fisher $p = 0.045$; Holm-adjusted $p = 0.182$). This pattern indicates a slight deviation from the economy-wide profile but no decisive Tertiary predominance.

The expectation that Financial Services exhibit a mixed profile with limited Primary (H1c) was evaluated next. The sector's 4×2 comparison to other industries was non-significant—$\chi^2(3) = 5.81$, $p = 0.121$, Cramér's $V = 0.240$—and no per-tier exact contrast remained significant after Holm correction. The data are consistent with a diversified tier distribution that is not statistically distinct from other sectors; therefore, H1c is rejected (no sectoral difference detected).

Potential performance differences across tiers within sectors were investigated using medians of turnover, EBITDA, and Growth YoY, tested with Kruskal–Wallis followed by Holm-adjusted Mann–Whitney comparisons and reporting Cliff's δ (H2). While some unadjusted pairwise contrasts approached conventional thresholds, effects were generally small-to-moderate and did not remain robust under multiplicity control. Given the small counts in Primary and Secondary, these null results may reflect limited power, yet the current evidence does not support systematic tier-linked performance advantages within sectors. While the vast majority of performance comparisons were non-significant after correction, Secondary-tier firms in the 'Other industries' group showed significantly higher EBITDA than their Tertiary counterparts, though this was an isolated finding. So, H2 is rejected at the 5% level after correction. (KW Within sectors and Pairwise MW tables in annex). The absence of robust short-term performance differentials across DC tiers should not be interpreted as evidence that higher data-capitalist intensity fails to create economic value. Rather, this pattern is consistent with a well-established body of literature on the productivity paradox of information technology, which documents delayed, non-linear, and often initially negative effects of digital investment on conventional performance metrics. Investments in data infrastructure, analytics platforms, organizational redesign, and specialized human capital are typically expensed upfront, depressing contemporaneous EBITDA, while their benefits materialize only with time, complementary process change, and learning.

Short-term competitive mobility was assessed via ΔRank = TOP 2023 − TOP 2024 (positive values indicate improvement / upward movement). No global differences across tiers were detected (Kruskal–Wallis $p \approx 0.75$), and pre-registered pairwise tests (Primary vs. NDC, Tertiary vs. NDC, Secondary vs. NDC) were non-significant after Holm adjustment, despite positive median ΔRank for some tiers. These results suggest that short-term rank volatility in the Latvian 'TOP 101' is not driven by data maturity. Given that strategic shifts into Data Capitalism require multi-year implementation cycles, a single-year rank change likely captures transient market factors—such as commodity price fluctuations or accounting adjustments—rather than the structural competitive advantages hypothesized in H3.Within the constraints of missing baseline ranks and small group sizes, there is no compelling evidence that higher tiers systematically drive short-run rank improvements, thus H3, as a measure of short-term competitive gain, is not supported within this one-year window.

While H2 and H3 are rejected at the 5% level, these null results provide a critical empirical snapshot. Rather than indicating a lack of value in data capitalism, the data may capture these firms during the 'trough' of the Productivity J-Curve. Primary and Secondary data capitalists in Latvia are likely engaged in the capital-intensive phase of data-centric transformation. In this stage, the costs of acquiring data, hiring specialized talent, and maintaining cloud infrastructure act as a drag on current EBITDA, effectively masking the long-term competitive advantages these firms are building.

Finally, the independence of DC_Tier from firm size and the size-robustness of the IT-Primary effect were examined (H4).

A contingency test of DC_Tier × turnover quartile was non-significant ($\chi^2 = 4.91$, df = 6, $p \approx 0.555$, $V \approx 0.20$), indicating no strong size–tier dependence. A Cochran–Mantel–Haenszel test pooling strata by turnover quartile indicates that the IT–Primary association persists after size control (MH common OR = 130.12; CMH $\chi^2(1) = 38.61$, $p = 5.18\times10^{-10}$), confirming that the IT-Primary overrepresentation persists across size strata.

The data analysis supports a strong and sector-specific statement: Primary data capitalism in Latvia is disproportionately concentrated in IT & Telecom, even after accounting for firm size.

Manufacturing & Consumer Trade exhibit only modest profile differences that do not condense into a simple Tertiary dominance once multiple testing is controlled, and Financial Services do not differ systematically from other industries in their tier composition. Within-sector performance contrasts by tier are weak to moderate and not consistently robust, and tier membership does not appear to drive short-run rank mobility in these data. These findings delimit where data-capitalist intensity is already central (IT & Telecom) and where its role is either diffuse or still emerging, thereby setting the stage for theory-driven explanations of sectoral pathways into data-centric business models and for targeted policy and managerial responses.

5. Conclusion

This study provides a granular, empirical map of data capitalism in Latvia. The findings demonstrate that data capitalism is not a uniform tide lifting all sectors equally but is a highly concentrated phenomenon. The investigation of the hypotheses yielded a nuanced picture of the Latvian corporate landscape. The stark overrepresentation of Primary data capitalists in IT & Telecom (H1a) stood in sharp contrast to the other sectors. The hypothesis that Manufacturing & Consumer Trade would be predominantly concentrated in non-Primary tiers (H1b) was not supported; a modest but not decisive deviation from the general industrial profile was observed. Similarly, no statistically significant difference in tier composition compared to other industries was found for the Financial Services sector, indicating a diversified but not distinctive profile (H1c).Significantly, no robust statistical evidence was found to support the common assumption that a higher DC_Tier directly translates into superior short-term financial performance (H2) or greater upward competitive mobility (H3) within this sample. The absence of a link between DC tiers and upward rank mobility (H3) highlights the temporal mismatch between data investment and market realization. Competitive mobility in a single-year window (2023–2024) appears to be a function of short-term volatility rather than a reflection of long-term strategic maturity. For 'Primary' data capitalists, the current lack of upward movement is consistent with the 'Productivity J-Curve' discussed earlier: firms at the forefront of the data economy may even experience temporary rank stagnation as they divert resources toward building intangible assets that have yet to translate into market-cap dominance.

This absence of clear, statistically significant findings may be partly attributable to limited statistical power; however, it challenges a simplistic narrative of data-driven dominance. It is suggested that the pathway from developing data-centric capabilities to achieving a measurable competitive advantage is complex, potentially non-linear, and may require more time to manifest in lagging performance indicators. The lack of a systematic relationship between DC tier and short-term financial performance or competitive mobility does not invalidate the economic significance of data capitalism. Instead, it points toward a non-linear, investment-intensive transition path. Consistent with the productivity paradox literature, firms at higher tiers—particularly Primary data capitalists—are likely incurring substantial upfront costs related to data infrastructure, organizational transformation, regulatory compliance, and experimentation. These costs may temporarily suppress observable profitability while building strategic assets whose value lies in future flexibility, scalability, and market positioning rather than immediate cash flows. Interpreted in this light, the findings suggest that data capitalism operates less as an instant performance enhancer and more as a long-horizon accumulation process. The Latvian evidence indicates that sectoral positioning determines where such investments are feasible and rational, while short-run accounting measures obscure their longer-term payoff structure.

The evidence strongly supports the conclusion that Primary data capitalism—where data itself is the core commercial asset—is overwhelmingly and robustly concentrated in the IT & Telecommunications sector. This sectoral distinction is not an artifact of firm size and suggests that sector-specific characteristics are critical determinants of a firm's pathway into the data economy. The absence of a clear link between DC tier and short-term performance also challenges a simplistic narrative of data-driven dominance, suggesting the pathway to competitive advantage is complex and potentially non-linear. From a policy perspective, our results caution against a "one-size-fits-all" approach to promoting the data economy; interventions should be tailored to the specific realities of each sector.
The study has several limitations. First, the DCMF was based entirely on publicly available information. This methodology, while transparent and replicable, may result in an underestimation of internal data strategies that are not disclosed for competitive reasons. Thus, the study's reliance on public disclosures introduces a potential 'visibility bias.' Firms in the IT & Telecommunications sector may be more vocal about their data capabilities because data is their primary product. In contrast, firms in more traditional sectors may treat advanced data analytics as a proprietary 'black box' or internal trade secret to maintain a competitive edge. While the DCMF attempted to capture these 'hidden' practices through the analysis of annual reports and recruitment data, the framework likely measures 'manifest' data capitalism more accurately than 'latent' or secretive internal optimizations. Future research involving internal audits or surveys could help uncover the extent of this hidden data-driven efficiency. This limitation could be overcome in future research by employing qualitative methods, such as in-depth case studies, to gain insight into proprietary data practices. Second, study is based on a snapshot from single point in time. The use of a one-year rank change (ΔRank) is a noisy proxy for competitive advantage, as strategic transformations like Data Capitalism often require a 3-to-5-year horizon to manifest in financial rankings. Short-term rank changes are highly susceptible to macro-economic noise—such as the 2023 energy price shocks or interest rate shifts—which may mask the emerging competitive signals of data-mature firms. A longitudinal study tracking this cohort over a longer duration is essential to separate transient rank volatility from structural, data-driven mobility. Third, the focus on the top 101 enterprises means the findings may not be generalizable to the small and medium-sized enterprises (SMEs) that constitute the backbone of the Latvian economy. A critical next step would be the replication of this study with a representative sample of SMEs. Finally, the results are specific to Latvia. Comparative studies in which the DCMF is applied to neighboring Baltic states or other similar small, open economies could reveal broader regional patterns and the influence of national industrial policy on the adoption of data-driven models.

References

Alshawawreh, A. R., Liébana-Cabanillas, F., & Blanco-Encomienda, F. J. (2024). Impact of big data analytics on telecom companies' competitive advantage. *Technology in Society*, *76*, 102459. https://doi.org/10.1016/j.techsoc.2024.102459

Arner, D. W., Barberis, J. N., & Buckley, R. P. (2015). *The Evolution of Fintech: A New Post-Crisis Paradigm?* (SSRN Scholarly Paper No. 2676553). Social Science Research Network. https://doi.org/10.2139/ssrn.2676553

Babina, T., Bahaj, S., Buchak, G., De Marco, F., Foulis, A., Gornall, W., Mazzola, F., & Yu, T. (2025). Customer data access and fintech entry: Early evidence from open banking. *Journal of Financial Economics*, *169*, 103950. https://doi.org/10.1016/j.jfineco.2024.103950

Banda, L., Mzyece, M., & Mekuria, F. (2022). 5G Business Models for Mobile Network Operators—A Survey. *IEEE Access*, *10*, 94851–94886. https://doi.org/10.1109/ACCESS.2022.3205011

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Bolin, G. (2022). The Value Dynamics of Data Capitalism: Cultural Production and Consumption in a Datafied World. In A. Hepp, J. Jarke, & L. Kramp (Eds.), *New Perspectives in Critical Data Studies: The Ambivalences of Data Power* (pp. 167–186). Springer International Publishing. https://doi.org/10.1007/978-3-030-96180-0_8

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence*. *The Quarterly Journal of Economics*, *117*(1), 339–376. https://doi.org/10.1162/003355302753399526

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Commun. ACM*, *36*(12), 66–77. https://doi.org/10.1145/163298.163309

Brynjolfsson, E., Rock, D., & Syverson, C. (2018). *The Productivity J-Curve: How Intangibles Complement General Purpose Technologies* (Working Paper No. 25148). National Bureau of Economic Research. https://doi.org/10.3386/w25148

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165–1188. https://doi.org/10.2307/41703503

Colangelo, G. (2024). Open Banking goes to Washington: Lessons from the EU on regulatory-driven data sharing regimes. *Computer Law & Security Review*, *54*, 106018. https://doi.org/10.1016/j.clsr.2024.106018

Gebauer, H., Arzt, A., Kohtamäki, M., Lamprecht, C., Parida, V., Witell, L., & Wortmann, F. (2020). How to convert digital offerings into revenue enhancement – Conceptualizing business model dynamics through explorative case studies. *Industrial Marketing Management*, *91*, 429–441. https://doi.org/10.1016/j.indmarman.2020.10.006

Kowalkowski, C., Windahl, C., Kindström, D., & Gebauer, H. (2015). What service transition? Rethinking established assumptions about manufacturers' service-led growth strategies. *Industrial Marketing Management*, *45*, 59–69. https://doi.org/10.1016/j.indmarman.2015.02.016

McAfee, A., & Brynjolfsson, E. (2012). Big Data: The Management Revolution. *Harvard Business Review*. https://hbr.org/2012/10/big-data-the-management-revolution, https://hbr.org/2012/10/big-data-the-management-revolution

McElheran, K., Li, J. F., Brynjolfsson, E., Kroff, Z., Dinlersoz, E., Foster, L., & Zolas, N. (2024). AI adoption in America: Who, what, and where. *Journal of Economics & Management Strategy*, *33*(2), 375–415. https://doi.org/10.1111/jems.12576

Mosch, P., Winkler, C., Eggert, C.-G., Schumann, J. H., Obermaier, R., & Ulaga, W. (2022). Driving or driven by others? A dynamic perspective on how data-driven start-ups strategize across different network roles in digitalized business networks. *Industrial Marketing Management*, *102*, 381–402. https://doi.org/10.1016/j.indmarman.2022.01.023

Mostaghel, R., Oghazi, P., Parida, V., & Sohrabpour, V. (2022). Digitalization driven retail business model innovation: Evaluation of past and avenues for future research trends. *Journal of Business Research*, *146*, 134–145. https://doi.org/10.1016/j.jbusres.2022.03.072

Oliveira, T., Thomas, M., & Espadanal, M. (2014). Assessing the determinants of cloud computing adoption: An analysis of the manufacturing and services sectors. *Information & Management*, *51*(5), 497–510. https://doi.org/10.1016/j.im.2014.03.006

Parker, G. G., Alstyne, M. W. V., & Choudary, S. P. (2017). *Platform Revolution: How Networked Markets Are Transforming the Economy—and How to Make Them Work for You*. W. W. Norton & Company.

Porter, M. E., & Heppelmann, J. E. (2014). How Smart, Connected Products Are Transforming Competition. *Harvard Business Review*. https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition, https://hbr.org/2014/11/how-smart-connected-products-are-transforming-competition

Ritala, P., Keränen, J., Fishburn, J., & Ruokonen, M. (2024). Selling and monetizing data in B2B markets: Four data-driven value propositions. *Technovation*, *130*, 102935. https://doi.org/10.1016/j.technovation.2023.102935

Shankar, V., Kalyanam, K., Setia, P., Golmohammadi, A., Tirunillai, S., Douglass, T., Hennessey, J., Bull, J. S., & Waddoups, R. (2021). How Technology is Changing Retail. *Journal of Retailing, Re-Strategizing Retailing in a Technology Based Era*, *97*(1), 13–27. https://doi.org/10.1016/j.jretai.2020.10.006

Stryker, C. (2025, July 8). *Was versteht man unter Human-in-the-Loop (HITL)?* https://www.ibm.com/de-de/think/topics/human-in-the-loop

Teece, D. J. (2018). Profiting from innovation in the digital economy: Enabling technologies, standards, and licensing models in the wireless world. *Research Policy*, *47*(8), 1367–1387. https://doi.org/10.1016/j.respol.2017.01.015