THE FIFTH WAVE

BRIE-ETLA Collection of Articles

Timo Seppälä | Tomasz Mucha | Juri Mattila (eds.)

THE FIFTH WAVE

BRIE-ETLA Collection of Articles

Timo Seppälä | Tomasz Mucha | Juri Mattila (eds.)

ETLA Economic Research
Elinkeinoelämän tutkimuslaitos – ETLA

Series B281
ISSN 0356-7443
ISBN 978-951-628-759-4 (printed)
ISBN 978-951-628-760-0 (pdf)

Taloustieto Oy, Helsinki 2023
Cover picture: www.shutterstock.com
Printing house: Next Print Oy

Suggested citation:
Seppälä Timo, **Mucha** Tomasz, **Mattila** Juri (Eds.) (2023).
The Fifth Wave – BRIE-ETLA Collection of Articles.
Helsinki: Taloustieto Oy.

Table of contents

INTRODUCTION TO COLLECTION OF ARTICLES

Beyond AI, Blockchain Systems, and Digital Platforms: Digitalization Unlocks Mass Hyper-Personalization and Mass Servitization 5

Article 1 Estimating Firm Digitalization: A Method for Disaggregating Sector-level Digital Intensity to Firm-level 25

Article 2 AI Diffusion Monitoring among S&P500 Companies: Empirical Results and Methodological Advancements 49

Article 3 What Have We Learned About Machine Learning? A Meta Analysis 81

Article 4 Blockchain-Based Deployment of Product-Centric Information Systems 99

Article 5 Digital Protocols as Accounting and Incentivization Mechanisms in Anti-Rival Systems: Developing a Shareable Non-Fungible Token (sNFT) 139

Article 6 The Little Engines That Could: Game Industry Platforms and the New Drivers of Digitalization 161

Article 7 Platform-Dependent Entrepreneurs: Power Asymmetries, Risks, and Strategies in the Platform Economy 173

Article 8 Supranationalism, Sino-American Technology Separation, and Semiconductors: First Observations 211

The Fifth Wave – BRIE-ETLA Collection of Articles

Tiivistelmä: Älykkäät teknologiat ovat viime vuosikymmenien aikana ohjanneet yritysten tuottavuutta, yhteiskunnan kehitystä ja talouskasvua. Jatkuva digitalisoituminen on mahdollistanut yksilöille ja organisaatioille älykkäiden työkalujen hyödyntämisen enenevässä määrin, mikä on johtanut aika ajoin häiritseviin muutoksiin markkinoilla ja sosiaalisissa sopimuksissa.

Älykkäiden teknologioiden ja yhteiskunnan huomion keskittyminen viimeisimpiin odotettuihin häiriöihin voi siirtyä nopeasti yhdestä teknologiailmiöstä toiseen. Esimerkiksi tekoälyä – vaikka se on edelleen kiistatta kuuma aihe tämän päivän keskustelussa – hädin tuskin julistettiin digitaalisen muutoksen kaikkein läpäisevimmäksi katalysaattoriksi seitsemän vuotta sitten. Samoin on oletettua, että seuraavan kolmen vuoden aikana uudet paradigman muutokset tapahtuvat häiritsevän teknologian kehityksen maisemassa, kuten tämän kirjan ensimmäisessä kappaleessa ennustamme.

Tutkimuksen mukaan koneoppimisen ja luonnollisten kielten menetelmien aikakausi tekoälyn häiriöissä on edelleen jatkumassa, ainakin toistaiseksi. Tosin tutkijat eivät ole kyenneet pääsemään yksimielisyyteen siitä, mikä teknologinen kehitys ja sen sovellukset tulevat nousemaan huomion keskipisteeksi nykykehityksessä. Lisäksi nykykehityksen kannalta välttämättömiä keskeisiä alustoja, liiketoimintamalleja, regulaatioita tai muita älykkäitä työkaluja ei ole vielä toistaiseksi tunnistettu.

Suurimmassa osassa nykyisistä yhteiskunnallisista ja organisaatioanalyyseistä käytetään kapeaa perspektiivinäkymää analysoimalla historiallista makro- tai mikrotaloudellista dataa. Uuden kehittyvän teknologian tapauksessa tarvitaan kuitenkin usein moniulotteista ja monitieteistä tutkimusta, jotta ymmärretään teknologian häiriöiden taustalla olevia monimutkaisia sosioekonomisia mekanismeja ja sitä, miten parhaiten voidaan navigoida yrityksiä ja maita teknologian aiheuttamien äärimmäisten turbulenssien aikana. Tähän Suomen tulisi tulevaisuudessa erityisesti innovaatio- ja teollisuuspolitiikassa panostaa. Nyt juoksemme pahasti jälkijunassa. Esimerkiksi ChatGPT löi meidät ällikällä, ja lähdemme tähänkin kehitysvaiheeseen pitkältä seitsemän vuoden takamatkalta.

Asiasanat: Digitaaliset alustat, Digitalisaatio, Lohkoketjujärjestelmät, Metaihminen-järjestelmät, Meta-organisaatiot, Operaatiot, Palveluiden tuottavuus, Palvelut, Palveluiden massapersonointi, Palvelullistuminen, Tekoäly, Tekoälyalustat, Tuottavuus, Älykkäät teknologiat

INTRODUCTION TO COLLECTION OF ARTICLES

Beyond AI, Blockchain Systems, and Digital Platforms:

Digitalization Unlocks Mass Hyper-Personalization and Mass Servitization

Timo Seppälä[a, b]
Tomasz Mucha[a]
Juri Mattila[c]

Affiliations:

[a] Aalto University, Finland
[b] ETLA Economic Research, Finland
[c] VTT Technical Research Centre of Finland Ltd, Finland

Acknowledgements:

We would like to thank all the institutions and the project that provided funding for this book and the underlying research work: Business Finland, the Berkeley Roundtable on the International Economy, ETLA Economic Research i.e., BRIE-ETLA research project, and Yrjö and Senja Koivunen's Foundation. We would also like to extend our gratitude to the Aalto University and their Department of Industrial Engineering and Management, especially to Risto Rajala. Additionally, we would like to express our appreciation towards our colleagues at ETLA Economic Research in regard to this book: Kimmo Aaltonen and Justiina Airas.

Abstract

The ever-progressing digitalization of the economy and society is unlocking new opportunities for organizations engaging in services. We are in the middle of a transformation of the service sector that can be likened to the advent of mass production in the 1940s. Based on recent advances and developments in artificial intelligence, digital platforms, and blockchain systems, we are witnessing the emergence of new digitalization phenomena of metahuman systems, artificial intelligence platforms, and meta-organizations. Jointly, these forces are shaping now, or will be in the near future, the service activities of organizations around the world. They enable mass hyper-personalized services and mass servitization – new types of high variety and high-volume service processes. Artificial intelligence applications like search and recommendation engines, and artificial intelligence platforms such as Google Maps, Chat GPT, BloombergGPT and Stable Diffusion can be perceived as early manifestations of the ongoing transformation. Already in the present day, applications and platforms such as these can be adopted in a wide range of downstream tasks, thus enabling personalized service experiences for audiences of one. While increasing the value of service offerings, mass hyper-personalization and mass servitization also have the potential to increase the productivity of service operations and the entire service sector, especially in the context of knowledge-intensive work. This introductory chapter gives us an opportunity to not only provide an overview of the articles included in this collection and their contributions, but also allows us to reflect and provide an up-to-date synthesis of key emerging concepts and research directions grounded in our research. Thus, this chapter in its own right goes above and beyond the articles included in this collection and contributes to the ongoing discussion on digitalization and the future of the service sector.

Keywords

Artificial intelligence, Artificial intelligence platforms, Blockchain systems, Digital platforms, Hyper-personalized services, Mass hyper-personalization, Mass services, Mass servitization, Metahuman systems, Meta-organizations, Operations, Productivity, Professional services, Service productivity, Service shops

Introduction

Mass hyper-personalization and mass servitization are grounded in the underlying megatrend of digitalization of business operations across industries. Digitalization is a multi-faceted and complex phenomenon that is continuously evolving (Calvino et al., 2018). Managers, entrepreneurs, policy makers, and consumers around the world are increasingly engaging with new forms of digitalization – a CEO consults BloombergGPT when preparing for a meeting with investors (Wu et. al., 2023), a shopkeeper in a country experiencing high inflation accepts payments in Bitcoin to hold value, a SaaS startup founder creates a service workflow as a distributed smart contract (Mattila, Hukkinen, Seppälä, 2017), the British Goverment drafts a risk-based framework for regulating the use of artificial intelligence (AI) (McCallum, 2023), a parent leaves their own car in a repair shop and rides Uber to pick up their kids from school. Overall, the new forms of digitalization and their applications are pervasive. They diffuse throughout the economy and society at a rapid pace and have a significant impact (Mucha & Seppala, 2022). As a result, the underlying changes require investigation of novel phenomena driving future productivity in the economy – this is especially true for service productivity.

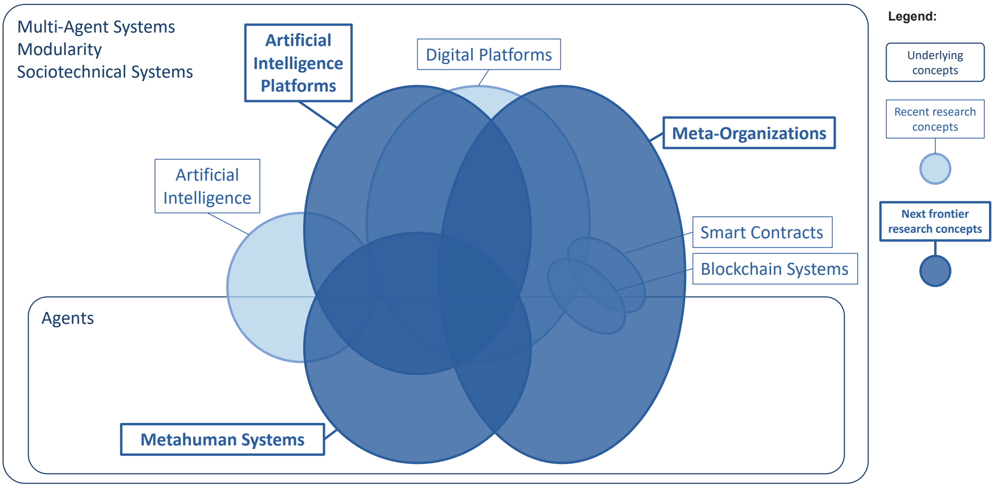
In terms of contributions towards Gross Domestic Product (GDP) and international trade, the importance of the service sector has increased, both nationally and globally (World Trade Organization, 2019). It is important to recognize, however, that “an economy’s prosperity does not depend on the relative size of its manufacturing or services sectors but on the productivity of the economy as a whole – which in turn depends on efficiencies and innovations across all sectors, and the extent to which they are mutually reinforcing” (World Trade Organization, 2019, p. 16). Hence, we need to understand the impact of the emerging digitalization trends on the service sector, servitization of industry, and their broader interconnections. Based on insights from the articles included in this collection, as well as synthesis of the recent digitalization research and our evaluation of the unfolding digitalization around the world we have identified mass hyper-personalization and mass servitization as vital emerging concepts driving future service sector competitiveness and productivity.

To understand the role of mass hyper-personalization and mass servitization, we need to consider them in the context of existing service processes. Three service process types are typically defined in research: professional services (high variety and low volume), service shops (medium variety and medium volume), and mass service (low variety and high volume) (Silvestro, Fitzgerald, Johnston & Voss, 1992; Silvestro, 1999). Confined by these definitions and the technological constraints of the past, relatively little attention has been dedicated to high variety and high-volume service provision. It is this type of service process, however, where we see the greatest untapped opportunity for improving service sector productivity in the future. We characterize **mass hyper-personalization** as an *efficient, dynamic, and high-volume process* *of targeting, designing, and delivering customized service experiences for an audience of one*, based on a set of unique criteria, e.g., by using foundation models that are prompted or fine-tuned with user-specific data. The path that is leading the service sector in the direction of mass hyper-personalization has been paved by the recent advances in digital technology – AI and blockchain systems – and the associated new ways of organizing economic activity – digital platforms and smart contracts. Servitization is typically depicted as a process of building revenue streams for company operations from services (Vandermerwe & Rada, 1988). Building on established conceptualizations of services, research on the digitalization of services has contributed to our understanding of how segmentation, customization, and servitization can impact firm productivity, and the process of how those changes take place. Digitalization enables the customization of services more productively than before (Marco, Vendrell-Herrero & Bustinza, 2018). However, the existing research on servitization has not sufficiently addressed the wide variety and high-volume aspects of digital services. While servitization has always included technological aspects, digital technologies have recently attracted increasing attention in this stream of literature (Van Ark, De Vries & Erumban, 2021; Linde, Frishammar & Parida, 2021). This has resulted in the recognizing of digital servitization where digital tools are the fundamental drivers in shifting a firm's business model from product-centric to service-centric (Kowalkowski et. al., 2017; Kohtamäki, Parida, Patel & Gebauer, 2020; Paschou, Rapaccini & Adrodegari, 2020). We predict that the next stage in digital servitization is **mass servitization**, which we define as *a universal high-volume transformation process of shifting from a product-centric business model to a service-centric approach by embedding learning, autonomy, and human interaction capabilities into emergent product-service bundles*.

This collection of articles presents insights on three inter-related themes of digitalization, which we consider essential in our quest to understand mass hyper-personalization and mass servitization as recently emerging aspects of digitalization. These themes are 1) AI, machine learning-based capabilities, and sociotechnical changes leading to the creation of metahuman systems in organizations (Mucha & Seppälä, 2020; Lyytinen, Nickerson & King, 2020; Dwivedi et.al., 2023); 2) blockchain-based systems and other intelligent tools underlying new types of distributed platforms or meta-organizations for collaboration (Lauslahti, Mattila & Seppälä, 2017; Hukkinen, Mattila & Seppälä, 2017; Mattila & Seppälä, 2018; Lumineau, Wang & Schilke, 2021); and 3) policy considerations for competition, innovation, digital technology stack, and platformized modes of operation (Cenamor & Frishammar, 2021, Holmström & Seppälä, 2020; Cutulo & Kenney, 2021).

To navigate and investigate this conceptually novel, evolving and intertwined terrain, we need to be armed with a vocabulary that allows us to capture and express what we encounter. Therefore, for the benefit of the reader, we collect and recap here some of the key concepts, which we first present as a carefully arranged visual map (Figure 1), and subsequently define in detail (Table 1). The list is not exhaus-

Figure 1 A mapping of selected key concepts related to digitalization, mass hyper-personalization, and servitization



tive and conceptual overlap is inevitable, because many of the definitions originate from distinct scholarly traditions or literatures and have different scope in terms of levels of analysis. We primarily draw these definitions from digital platforms, blockchain, and artificial intelligence literatures and, when needed, we refer to economics, information systems, and other disciplines of research. While this arsenal clearly reflects the complexity and multi-faceted nature of digitalization, it allows us to identify areas of future focus for scholars, business practitioners, and policy makers.

Hierarchy/arrangement of selected concepts/definitions:

- Underlying concepts:
  - Agents
  - Modularity
  - Multi-Agent Systems
  - Sociotechnical Systems
- Key concepts in recent research on digitalization:
  - Digital Platforms
  - Artificial Intelligence
  - Smart Contracts
  - Blockchain Systems
- Emerging concepts shaping future research on digitalization:
  - Metahuman systems
  - Artificial Intelligence Platforms
  - Meta-Organizations

Table 1 Selected definitions of key concepts

	Definition	Source
Underlying concepts:		
Agent	Is an individual human, but also in some settings an information systems artifact or an organization, possessing "the ability to accept rights and responsibilities for ambiguous tasks and outcomes under uncertainty and to decide and act autonomously."	Baird & Maruping, 2021; Lyytinen & Newman, 2008
Modularity	Is an approach where different parts of the product and/or service and/or software are designed and manufactured by separate, specialized working groups working independently of one another. The "modules" could then be connected and (in theory at least) would function seamlessly, if they as they confronted to a predetermined set of design and manufacturing rules. With modularity enforced, it is possible to change pieces of the system without redoing it whole. Designs and manufacturing become flexible and capable of evolving at the module and system levels.	Baldwin & Clark, 2000
Multi-agent systems	Consist of autonomous entities know as agents. Agents collaboratively solve tasks, yet they offer more flexibility due to their inherent ability to learn and make autonomous decisions. Agents use their interactions with neighboring agents or with the environment to learn new context and actions. Subsequently, agents use their knowledge to decide and perform an action on the environment to solve their allocated task.	Dorri, Kanhare, Jurdak, 2018
Sociotechnical systems	Are "any organizational system viewed as a multivariate system consisting of four interacting and aligned components – task, structure, actor, and technology."	Lyytinen & Newman, 2008
Key concepts in recent research on digitalization:		
Digital Platforms	Are an evolving organizations and meta-organizations that: 1) federate and coordinate constitutive agents who can innovate and compete; 2) create value by generating and harnessing economies of scope in supply and/or in demand side of the markets; and 3) entail modular technological architecture composed of the core and periphery.	Gawer, 2014;
Artificial Intelligence	Is the frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems. It is, furthermore, multidimensional and can be presently viewed from the following perspectives: 1) Data analytics; 2) Sensing and situation awareness; 3) Natural language and cognition; 4) Interaction with humans; 5) Digital skills, interactions in work life; 6) Machine learning; 7) System level and systemic impact; 8) Computing equipment, platforms, services and ecosystems; 9) Robotics and machine automation – the physical dimension of AI; 10) Ethics, moral, regulation and legislation.	Berente et al., 2021; Ailisto et al., 2018
Smart Contracts	Are digital computer programs that: 1) are written in computer code and formulated using programming languages; 2) are stored, executed and enforced by a distributed and replicated blockchain network; 3) can receive, store and transfer digital assets of value; and 4) can execute with varying outcomes according to their specific internal logic.	Lauslahti, Mattila & Hukkinen, Seppälä, 2018
Blockchain systems	Are 1) open source and open access technology compositions; 2) comprising non-hierarchal peer-to-peer networks without any single point of failure or control; 3) which maintain consensus over cryptographically concatenated, shared, replicated append-only data structures; 4) according to deterministic self-contained consensus algorithms, void of external inputs such as validation by central authorities or off-chain signaling.	Mattila, 2021
Emerging concepts shaping future research on digitalization:		
Metahumans systems	Are new, emergent, sociotechnical systems where machines that learn join human learning and create original systemic capabilities.	Lyytinen, Nickerson & King, 2020
Artificial Intelligence Platforms	Are digital platforms which critically rely on AI technologies in at least one of the following areas: 1) federation and coordination of constitutive agents; 2) value creation; or 3) technological architecture.	Mucha & Seppälä, 2020
Meta-organizations	Are cross-organizational systems where multiple agents (human, metahuman system, and legally autonomous organization) interact in a 1) dynamic, 2) interoperable, 3) intelligent, 4) federated, and 5) coordinated manner, thus enabling them to create unique and context specific bundles of product-service design and delivery.	This article

We grouped the key concepts presented here into three sets, which reflect the chronological and conceptual progression of academic research and evolution of digitalization. The underlying concepts are the broadest and most seasoned ones. Apart from connecting our work to long-established research, they also show how our understanding needs to be periodically revised, as technology and society advance. For example, we used to consider only humans or organizations as agents. Now, however, technology artifacts have been endowed with much higher levels of autonomy and capabilities, thus exhibiting agentic properties (Baird & Maruping, 2021). Concepts related to the recent research on digitalization are central to understanding the articles included in this collection, which we will present next. Finally, the emerging concepts reflect our newly informed insights, which are based on our synthesis of findings and contributions from the articles included in this collection, recent literature on digitalization, and active engagement with digitalization taking place in the industry and society.

Contributions in this collection of articles

This collection of articles introduces three sets of themes. The first theme, articles one to three, describe when and how companies have started to adopt AI leading to creation of metahuman systems in organizations. The second theme, articles four to six, explain how blockchain systems have been considered by companies and new distributed collaborative meta-organizations. The third theme, articles seven to eight, consider policy implications for competition, innovation and industries primarily in the context of technology stack and digital platforms.

The first three articles in this collection (Mucha & Seppälä, 2021; Mucha & Seppälä, 2022; Mucha, Seppälä & Gustafsson, 2023) examine the technology diffusion and corporate adaptation of artificial intelligence technologies and the increasing importance of AI platforms. The first article proposes a method for estimating firm-level digital intensity based on industry sector level data, which can be used to understand firm digitalization among its peer group. The proposed method considers firms' participation in multiple industries, uses reference sector-level digital intensity scores, and is replicable and reproducible. The second article proposed a method for monitoring the commercial diffusion of technology which captures the temporal progression of technology adoption by organizations and relies on qualitative content coding. It provides transparent, replicable, updatable, and granular results that are illustrated using the case of AI diffusion among S&P 500 companies. The third article takes a sociotechnical system perspective on the micro foundations of capabilities and develops an integrative conceptual framework to extend understanding of organizational capabilities in the context of machine learning (ML) initiatives. The framework incorporates a temporal dimension, and multiple propositions are developed using anecdotal evidence.

Three contributions, article four (Mattila, Seppälä, Valkama, Hukkinen, Främling, Holmström, 2021), article five (Hakanen, Eloranta, Marttila & Amadae, 2023) and article six (Mattila, Seppälä & Salakka, 2021) of this collection of articles discuss new blockchain systems and other intelligent tools and their impacts on organizations and markets. The fourth article proposes a blockchain-based approach for product information management, which aims to collect product life-cycle data, maintain an accurate single state of product information, and provide economic incentives for solution deployment. The evaluation identifies challenges in deploying blockchain-based solutions in the current industrial landscape, but the paper lays the foundation for a self-sustained and self-incentivized deployment approach. The fifth article talks about other kinds of blockchain systems i.e., distributed ledger technologies (DLT), primarily designed to facilitate the exchange of unique, scarce items. This paper presents an alternative decentralization protocol based on anti-rival goods. The authors explain the technical approach behind the concept, referred to as shareable non-fungible tokens (sNFTs), and illustrate their argumentation by presenting a decentralized platform for sharing and streaming data. The sixth article considers the game industry's expertise in building virtual economies that can establish data product markets, potentially challenging digital platform incumbents. To protect the Finnish game industry and economy, policymakers should understand the resources, protocols, and regulative frameworks required to foster new businesses and industrial growth in new digital infrastructures.

Two contributions, article seven and eight of this collection of articles discuss competition, innovation and industry policy implication (Cutulo & Kenney, 2021, Holmström & Seppälä, 2020). The seventh article discusses the significance of digital platforms, especially the power asymmetry between platforms and ecosystem members is intrinsic to their economics and technological architecture. Article seven suggests that entrepreneurs in the platform ecosystem are more usefully termed "platform-dependent entrepreneurs" (PDEs) and explores strategies to mitigate their dependence. Additionally, the article provides a framework for policy makers to consider regulating platform-organized markets. The eight article focuses on the US-China trade conflict and the potential technology separation that could disrupt global value chains of digital technologies, particularly in the lower hardware levels of the technology stack. The article highlights the potential implications for Europe and smaller open economies such as Finland and explores different options for Europe if the technological separation continues.

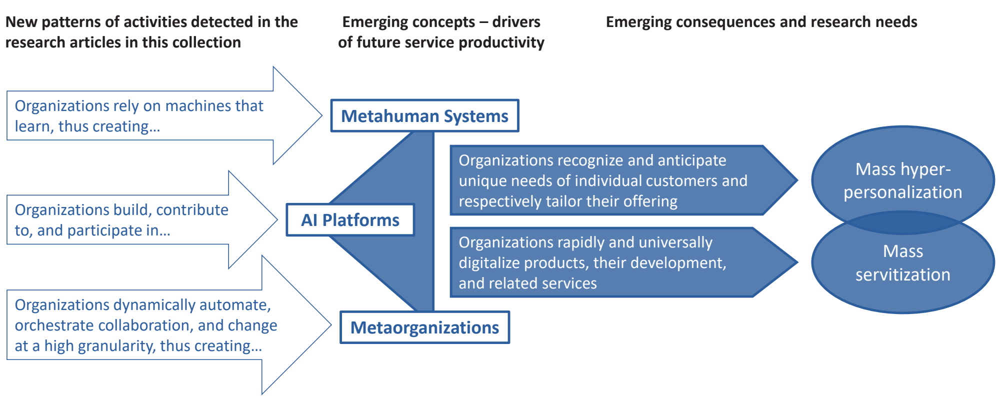
Jointly, the three themes addressed by articles in this collection indicate the directions in which digitalization of industry and society is inevitably evolving. This direction, in our view, is hyper-personalization of services and mass servitization. These two phenomena are grounded in the emergence of metahuman systems, AI platforms, and meta-organizations. The articles in this collection identified and explored the harbingers of these nascent systems or their building blocks. Based on the early evidence, already at this stage, we recognize the transformative impact of these systems on various industries. In the upcoming section, we take a more in-depth look on these future research areas, and thereafter we develop a research agenda focused on the implications for hyper-personalization and mass servitization. This is an area where we anticipate the impact of these systems will be particularly significant.

The emerging cornerstones of mass hyper-personalization and mass servitization: Metahuman systems, AI platforms, and meta-organizations

Understanding key technologies and their potential impacts is merely a starting point. Ultimately, technologies do not determine outcomes — people, organizations, and institutions interacting with technologies do (Emery, 1993; Leavitt, 1965). In short, technology enables action. It is ours to decide how to apply it, and with what kinds of consequences. Vice versa, technology deployment and its context are influenced by strategies, regulation, and policies. Therefore, we need to better understand the broader sociotechnical aspects of the emerging drivers or cornerstones of operations and service productivity.

By building on the insights from the articles included in this collection and complementing these with our readings of the recent literature on digitalization, as well as our perception of the unfolding digitalization around the world, we identify metahuman systems, AI platforms, and meta-organizations as the emerging concepts shaping future research on digitalization, particularly in relation to mass hyper-personalization and mass servitization (Figure 2).

Figure 2 The emerging patterns of activities and concepts shaping future research on digitalization



Metahuman systems are new, emergent, sociotechnical systems where machines that learn join human learning and create original systemic capabilities at the level of teams or, potentially, organizations (Lyytinen et al., 2020; Mucha et al., 2022, Forthcoming). These new capabilities are distinct, because without ML technologies that are learning and adapting it would not be practically or technically feasible to reach the required levels of performance within these systems (Mucha et al., 2023). Metahuman systems will impact operations and service delivery from the perspective of both the organizations providing the service as well as those of customers receiving, co-creating or co-operating within the service.

The distinction between internal and external impact of metahuman systems on organizations is important because it highlights the sweeping impact of metahuman systems on operations and service productivity. First, many organizations internally consider knowledge workers as providers of internal services to other units, functions, or roles (Davis, 1996). Machines that learn already now can or soon will be able to keep track of sets of actions of individual employees and in conjunction with that start modifying own behavior to increase the level of personalization for the need of these employees. If successfully executed and developed into hyper-personalization, these metahuman systems will potentially improve the baseline performance by, for example, lowering variance, increasing throughput, or improving output quality (Mucha et al., 2023). Clearly, some metahuman systems will also be re-imagined and novel, rather than incrementally developed versions of the preceding sociotechnical systems (Mucha et al., 2021, Forthcoming). However, even more impactful productivity gains can be reaped by organizations leveraging metahuman systems to render services that are more valuable than the status quo and serve external customers. By creating offering that is better tailored to external customer needs, especially those “jobs to be done” that are unique, important, and insufficiently catered to. Thus, metahuman systems will, in many cases, form the fundamental building blocks underlying hyper-personalization.

This line of reasoning is also salient to understanding the role of metahuman systems in enabling and fostering mass servitization. One of the stumbling blocks on the transformation path from product to service logic is the scalability of human resources and the ability to respond to unique customer needs (Zhang & Banerji, 2017). These challenges in our view have prevented, thus far, servitization from happening on a mass scale. Metahuman systems, however, will help organizations to scale human knowledge and capacities better by encapsulating some of these into technology that is essentially freely scalable (Mucha et al., 2021).

Another cornerstone of future operations and service productivity is the increasingly critical role of AI in the functioning of digital platforms, thus the emergence of AI platforms (Mucha & Seppala, 2020). While the platformization of the economy is already a well-established trend, we have seen only very preliminary impacts of AI in this domain compared to what is already now feasible from a technology viewpoint.

AI platforms provide a backbone to many individual organizations actively leveraging or constructing metahuman systems (Mucha & Seppala, 2020). Consequently, understanding the role of AI platforms in this capacity will be pivotal.

We, furthermore, need to consider both innovation platforms and transaction platforms having both important and unique own contribution to this evolution (Cusumano et al., 2020). Innovation platforms will be both fostering and constraining some organizational uses of AI. This will have an important impact on the competitive dynamics of service sector, because uneven access, maturity of, or ability to leverage AI will partially determine the outcomes of mass hyperpersonalization efforts of organizations (Mucha & Seppala, 2020). Transaction platforms, on the other hand, will play a crucial role in distributing and disseminating services or information about these services. Furthermore, transaction platforms might constitute some of the marketplaces where critical enablers of hyper-personalization will be exchanged. This logic extends to mass servitization, because of the constant pressure and efforts towards platformization of industrial sector. Here, it is important to recognize the role of newly emerging AI platforms (start-ups), which are distinct from hyperscalers (Mucha & Seppala, 2020). These AI platforms will likely play an important role in mass servitization because their offering might be centered around specific servitization use cases.

Finally, we recognize that meta-organizations emerge as the third novel cornerstone of future productivity growth in the service sector. While past research has already identified the concept of meta-organizations seen as organizations comprising multiple legally autonomous entities (Gawer, 2014; Gulati et al., 2012), our conceptualization updates that definition to reflect multi-level interactions of various agents constituting meta-organizations. We propose to include within the scope of meta-organizations other types of agents as well – individual humans and metahuman systems emerging within organizations. Furthermore, these agents must be able to interact in a 1) dynamic, 2) interoperable, 3) intelligent, 4) federated, and 5) coordinated manner, which enables them to create unique and context specific bundles of product-service design and delivery. Thus, the interactions constituting the fabric of meta-organizations are present not only within a single organization, but also might frequently cross the organizational boundaries.

AI platforms represent one type of meta-organization, but the scope of meta-organization as a concept is nevertheless much broader than that. For example, the interactions between the actors might be governed by a smart contract and not necessarily rely on a digital platform logic. To add to that, multi-level aspect of meta-organizations reveals important contributions of individual humans and metahuman systems to render product-service bundles. For example, human annotators who label training data for ML models play an important role from the perspective of the system as a whole. Equally, ML-based digital artifacts might drive and constrain actions of human actors or even entire organizations. Thus, meta-organizations constitute a distinct and complementary determinant of mass hyper-personalization and mass servitization.

A tentative research agenda

We believe that with the increasing computing power, the proliferation of AI to firms and digital platforms, and the related emergence of new organizing logics, we are amid a service sector transformation resembling the advent of mass production in the 1940s. The resulting service sector productivity dynamics will be driven by mass hyper-personalization and mass servitization. Our collection of articles points to several fruitful areas for future research inquiry in services, servitization and productivity in the context of AI platforms, metahuman systems, and meta-organizations.

The tentative nature of the proposed agenda reflects the nascent stage of the phenomena we urge scholars to study. Furthermore, in outlining questions for future research, we primarily concentrate on aspects that drive nuanced understanding and contextually rich micro-level perspective. This reflects the complex sociotechnical dimension that we need to understand to better appreciate often nuanced differences between traditional information technology and AI. Therefore, the proposed research directions concentrate on in-depth (case) studies that provide understanding of novel phenomena. However, we expect that macro-perspective approach will soon become viable as well, given the rapidity and pervasiveness of changes that take place in, at least, some of the relevant areas. For instance, in January 2023 ChatGPT became the fastest growing consumer application ever, beating even digital platforms such as TikTok or Instagram (Hu, 2023). Thus, research utilizing quantitative data will need to be developed as well.

1 Metahuman systems – Foundation models, operations, and service automation

By harnessing modern computing resources, abundant data, and continuously advancing algorithms in operations and services, we have greatly improved state-of-the-art computer performance on many tasks such as speech recognition, image recognition, and generation of text, audio, and images. Some of these capabilities have been packaged in the form of foundation models, which have been trained on broad data, can be further fine-tuned to specific tasks and recombined to create new intelligent tools such as ChatGPT and GPT-4.

These novel technologies have the potential to change the ways modern organizations work – the roles of people, the routines they enact, the products and services they deliver, and productivity they achieve. This transformation, however, is not merely about technological progress. Productively integrating these intelligent tools into mass hyper-personalization and servitization of industry requires that we explore and understand the new opportunities and limits of digital automation. Particularly, the understanding of digital automation limits remains downplayed and overlooked.

The naïve view is that simply with more data and computing resources the performance of these new forms of digital automation increases. However, for private sector companies to leverage these tools and drive productivity improvement, as well as for innovation and growth policy actions to foster that development, we need a more in-depth understanding and paradigmatic case examples of the newly redefined constraints of digital automation.

Thus, we propose the following research questions to drive research along this dimension.

Research question 1.1:	What is the foundation model application landscape within and outside of generative AI applications for product and service companies?
Research question 1.2:	What do mass hyper-personalized service and mass servitized engagements and experiences mean for knowledge work and worker?
Research question 1.3:	What are the limits to mass hyper-personalization and mass servitization in metahuman systems? How companies drive productivity within these boundaries?
Research question 1.4:	What are the limits to productivity improvement in organizations relying on foundation models, other types of machine learning, or metahuman systems? How do companies drive productivity within these boundaries?

2 Artificial intelligence platforms and service firm productivity

The scale of artificial intelligence (AI) platforms, their workloads, and range of offering have been increasing continuously. In the early days of AI use by digital platforms, these technologies were just one of their tools in the toolbox and were utilized predominantly in internal processes. It is important to recognize that subsequently many digital platform companies have not only invested in research and development of AI for improving their own operations but have also looked for the ways to productize AI applications and create own AI ecosystems. The resulting universe of AI platforms has been further enriched by rapid proliferation of various AI services and emerging AI platforms targeting specific services, industries, or market segments.

This growth creates numerous opportunities for service firms, but it is also full of challenges. The barriers to accessing state-of-the-art AI in the form of the latest machine learning models and particularly foundation models are disappearing. This is illustrated by Microsoft making its search engine become more conversational and Amazon partnering with Hugging Face to enable easy fine-tuning and deployment of latest models. This deceptive ease is coupled with many open issues regarding explainability, ownership, and legal basis to name just a few. To further complement the picture, various organizations including government agencies and non-profits are also experimenting with and leveraging these new AI tools. The resulting dynamics and the pivotal role of AI platforms is neither explored by scholars nor well-recognized by practitioners concerned with improving productivity of service firms.

Thus, we propose the following research questions to trace the development of the AI platform as a central feature of the contemporary digital economy and consider the consequences from the perspective of productivity and innovation policy.

Research question 2.1: What are the implications of artificial intelligence platforms integration and interoperability to company product and service portfolio management?

Research question 2.2: How are mass hyper-personalization and mass servitization designed, delivered, and organized by firms participating in AI platforms?

Research question 2.3: How is the productivity of service firms impacted by their participation in artificial intelligence platforms?

Research question 2.4: How is the productivity of service firms impacted by government and non-profit participation in artificial intelligence platforms?

3 Meta-Organizations – The new system architectures for productivity in operations, services, mass hyper-personalization and mass servitization

As various IT systems are becoming increasingly integrated to one another because of digitalization, entirely new modes of mass hyper-personalization and mass servitization are enabled through product, service, and process automation. As manufacturing transitions from a product-model-centric philosophy to a more object-oriented paradigm, product individuals become actors that can be tracked, mass customized and hyper-personalized dynamically over their entire life cycles in unprecedented ways. As product individuals are transformed into personalized service actors with individuality and embedded intelligence, new types of meta-organizations emerge where humans and product systems dynamically interact in mass-servitized and hyper-personalized manner uniquely according to every specific situation.

Simultaneously, in a similar trend of development, new types of platform innovations are enabling more individual user-oriented service logics in digital platforms. For example, through blockchain-based smart contracting platforms, digital workflow processes can be individually tailored, mass servitized, and hyper-personalized in entirely novel and democratized ways. Due to the decentralized nature of such systems, genuine switch-role markets can be generated in a new manner that enables much more dynamic modes of interaction between actors in meta-organizational structures.

Scholarly work falling under this topic should output paradigmatic case examples based on research engaged with practice. This will likely require concentrating on individual sectors, industries, or businesses to surface high-granularity data. Overall, developing insights into the new systems architectures and their impact on productivity will be one of the key objectives of this future research. Therefore, we propose the following research questions.

Research question 3.1:	What are the new micro-modular e.g., foundation model based and other, system architectures of service and servitized product firms successfully employing digital automation?
Research question 3.2:	What are the implications of these new service system architectures to global value chains?
Research question 3.3:	What are the implications of these new service systems architectures to productivity and what role do mass-personalization and mass servitization play in that?
Research question 3.4:	What are the innovation, industry and competition policy implications of these new service systems architectures?

Concluding remarks

Metahuman systems, artificial intelligence platforms, and meta-organizations are likely to continue affecting how work, especially knowledge work, is done. These digitalization phenomena converge to enable design and delivery of mass hyper-personalization of services and mass servitization, thus impacting how value is created and captured by companies representing the majority of the economy. The difficulty of predicting how these will affect different industries is due in part to their pervasive impacts. As mentioned in multiple articles in this collection, the application of these novel technologies is often characterized by their ubiquitous, persistent, and deep integration with other forms of economic activity. The initial applications are often generative, thus sparking further innovation which makes predicting the future difficult.

We believe that metahuman systems, artificial intelligence platforms, and meta-organizations are likely to be powerful organizing principles for companies and other organizations, for industries, the economy, and society over the coming years. Scholars interested in contemporary organizations and industries, or innovation and competition must consider how metahuman systems, artificial intelligence platforms and meta-organizations facilitate, constrain, channel, and change economic or social activity.

We anticipate a rising new "TIDE" of further studies related to metahuman systems, artificial intelligence platforms, and meta-organizations towards mass hyper-personalized operations and service design, delivery and experiences.

References

Ailisto, H., Heikkilä, E., Helaakoski, H., Neuvonen, A. & Seppälä, T. (2018). Tools for Artificial Intelligence Discussion. Article Series of Government's analysis, assessment and reserch activities No. 27.

Baird & Maruping. (2021). The next generation of reserch on IS use: A theoretical framework of delegation to and fromagentic artifacts. MIS Quaterly. Vol 45., Issue 1, pp. 315-341.

Baldwin, C. & Clark, K. (2000). Design Rules: The Power of Modularity. The MIT Press, Cambridge, Massachusetts, USA.

Berente, N., Gu, B., Reckar, J., & Santhaman, R. (2021). Managing Artificial Intelligence. MIS Quaterly. Vol. 45. Issue 3. pp. 1433-1450.

Calvino, F., Criscuolo, C., Marcolin, L., & Squicciarini, M. (2018). *A taxonomy of digital intensive sectors*. https://doi.org/10.1787/f404736a-en

Cenamor & Frishammar. (2021). Openness in platform ecosystems: Innovation strategies for complementary products. Research Policy. Vol. 50, Issue 1, 104148

Cusumano, M. A., Yoffie, D. B., & Gawer, A. (2020). The Future of Platforms. *MIT Sloan Management Review*, *61*(3), 46–54.

Cutulo & Kenney, M. (2021). Platform-Dependent Entrepreneurs: Power Asymmetries, Risks and Strategies in the Platform Economy, Academy of Management Perspectives, Vol. 35. No. 4, https://doi.org/10.5465/amp.2019.0103

Davis, T. R. V. (1996). Managing Knowledge-Work Support Functions. *Journal of General Management*, *22*(1), 68–86. https://doi.org/10.1177/030630709602200105

Dorri, A., Kanhare, S., & Jurdak, R. (2018). Multi-Agent Systems: A Survey. *IEEE Access*, vol. 6, pp. 28573-28593.

Dwivedi, Y., et. al. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management. Vol. 71, 102642.

Emery, F. (1993). Characteristics of Socio-Technical Systems. In E. Trist, H. Murray, & B. Trist (Eds.), *The Social Engagement of Social Science, Volume 2* (pp. 157–186). University of Pennsylvania Press; JSTOR. http://www.jstor.org.libproxy.aalto.fi/stable/j.ctt1bj4q98.10

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, *43*(7), 1239–1249.

Gulati, R., Puranam, P., & Tushman, M. (2012). Meta-organization design: Rethinking design in interorganizational and community contexts. *Strategic Management Journal*, *33*(6), 571–586.

Hakanen, E., Eloranta, V., Marttila., & Amadae, (2023). Digital protocols as accounting and incentivization mechanisms in anti-rival systems – Developing a shareable non-fungible token (sNFT). In Seppälä, T., Mucha, T., & Mattila, J. (eds) The Fifth Wave: BRIE-ETLA Collection of Articles, Taloustieto Oy, Helsinki.

Holmström, H. & Seppälä, T. (2020). Supranationalism, Sino-American Technology Separation, and Semiconductors: First Observations, ETLA Working Paper No. 82.

Hu, K. (2023, February 2). ChatGPT sets record for fastest-growing user base – Analyst note. *Reuters*. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

Hukkinen, T., Mattila, J. & Seppälä, T. (2017). Distributed Workflow Management with Smart Contracts. ETLA Reports No. 78.

Kohtamäki, Parida, Patel & Gebauer, (2020). The relationship between digitalization and servitization: The role of servitization in capturing the financial potential of digitalization. Technological Forcasting and Social Change. Vol. 151, 119804.

Kowalskowski, C., Gebauers, H., Kamp, B., & Parry, G. (2017). Servitization and deservitization: Overview, concepts, and definitions; Industrial Marketing Management, Vol. 60 pp. 4-10.

Lauslahti, K., Mattila, J. & Seppälä, T. (2017). Smart Contracts - How will Blockchain Tecnology Affect Contractual Practices? ETLA Reports No. 68

Lauslahti, K., Mattila, J., Hukkinen, T. & Seppälä, T. (2018) Expanding the Platform: Smart Contracts as Boundary Resources. In Sedlund, A., Lindblom, A., Mitronen, L. (eds) Collaborative Value Co-creation in the Platform Economy. Translational System Sciences, Voll 11. Springer, Singapore.

Leavitt, H. J. (1965). Applied organizational change in industry, structural, technological and humanistic approaches. *Handbook of Organizations*, *264*.

Linde, L., Frishammar, J., & Parida, V. (2021). Revenue models for digital servitization: a value capture framework for designing, developing, and scaling digital services. IEEE Transactions on Engineering Management, iss. 70, no. 1, pp. 82-97.

Lumineau, Wang & Schilkke, (2021). Blockchain Governance - A New Way of Organizing Collaborations? Organizational Science. Vol. 32, Iss. 2, pp. 500-521.

Lyytinen, K. & Newman, M. (2008). Explaining information systems change: a punctuated socio-technical change model. European Journal of Information Systems. Vol. 17, pp. 589-613.

Lyytinen, K., Nickerson, J. V., & King, J. L. (2020). Metahuman systems = humans + machines that learn. *Journal of Information Technology*, 0268396220915917. https://doi.org/10.1177/0268396220915917

Marco, O-B., Vendrell-Herrero, F., & Bustinza, O. (2018). Uncovering Productivity Gains of Digital and Green Servitization: Implications from the Automotive Industry. Sustainability 10, no. 5: 1524.

Mattila, J. (2021). Blockchain Systems as Multi-seded Platforms. ETLA Economic Research Series A, No. 51.

Mattila, J. & Seppälä, T. (2018). Distributed Governance in Multi-sided Platforms: A Conceptual Framework from Case Bitcoin. In Sedlund, A., Lindblom, A., Mitronen, L. (eds) Collaborative Value Co-creation in the Platform Economy. Translational System Sciences, Voll 11. Springer, Singapore.

Mattila, J., Seppälä, T. & Salakka, K. (2021). The Little Engines that Could - Game Industry Platforms ad the New Drivers of Digitalization,. ETLA Brief No. 101.

Mattila, J. & Seppälä, T., Valkama, P., Hukkinen, T., Främling, G. & Holmström, J. (2021). Blockchain-based Deployment of Product-centric Information Management Systems, Computer in Industry, Vol. 125 (2), pp. 1-14.

McCallum, S. (2023). UK rules out new AI regulator. https://www.bbc.com/news/technology-65102210 (retrieved 3.4.2023).

Mucha, T., Ma, S., & Abhari, K. (2022). Beyond MLOps: The Lifecycle of Machine Learning-based Solutions. *AMCIS 2022 Proceedings*, 11. https://aisel.aisnet.org/amcis2022/sig_adit/sig_adit/9

Mucha, T., Ma, S., & Abhari, K. (Forthcoming). Riding a Bicycle While Building Its Wheels: The Process of Machine Learning-Based Capability Development and IT-Business Alignment Practices. *Internet Research*.

Mucha, T., Seppälä, J., & Puraskivi, H. (2023). Context Changes and the Performance of a Learning Human-in-the-loop System: A Case Study of Automatic Speech Recognition Use in Medical Transcription. *Proceedings of the 56th Hawaii International Conference on System Sciences*. https://hdl.handle.net/10125/103014

Mucha, T., & Seppala, T. (2020). *Artificial Intelligence Platforms – A New Research Agenda for Digital Platform Economy* (SSRN Scholarly Paper ID 3532937). Social Science Research Network. https://doi.org/10.2139/ssrn.3532937

Mucha, T., & Seppälä, T. (2021). Estimating Firm Digitalization: A Method for Disaggregating Sector-Level Digital Intensity for Firm-level. MethodsX, Vol 8, 10233

Mucha, T., & Seppala, T. (2022). AI Diffusion Monitoring among S&P500 Companies: Empirical Results and Methodological Advancements. *DIGIT 2022 Proceedings*. https://aisel.aisnet.org/digit2022/13

Mucha, T., Seppälä, T., & Gustafsson, R. (2021). *Yksitoista oppia koneoppimisen meta-analyysistä* (Policy Brief No. 4/2021; Business Finland Policy Brief). Business Finland. https://www.etla.fi/julkaisut/yksitoista-oppia-koneoppimisen-metaanalyysista/

Paschou, Rapaccini & Adrodegari, (2020). Digital servitization in manufacturing: A systematic literature review and research agenda. industrial Marketing Management. Vol 89., pp. 278-292.

Silvestro (1999), Positioning services along the volume-variety diagonal: The contingencies of service design, control and improvement. International Journal of Operations & Production Management. Vol. 19., No. 4, pp. 399-421.

Silvestro, Fitzgerald, Johnston & Voss. (1992), Towards a Classification of Service Processes. International Journal of Service Industry Management, Vol. 3., No. 3.

Van Ark, B., De Vries, K., & Erumban, A., (2021). How to not miss a productivity revival once again. National Institute Economic Review, no. 255, pp. 9-24.

Vandermerve & Rada. (1988), Servitization of business: Adding value by adding services. European Management Journal. Vol. 6, Iss. 4, pp. 314-324.

World Trade Organization. (2019). World Trade Report 2019. The Future of Services Trade. WTO Secretariat Geneva.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D. & Mann, G. (2023). BloombergGPT: A Large Language Model for Finance. arXiv:2303.17564 (cs).

Zhang, W., & Banerji, S. (2017). Challenges of servitization: A systematic literature review. Industrial Marketing Management, 65, 217–227. https://doi.org/10.1016/j.indmarman.2017.06.003

ARTICLE 1

Estimating Firm Digitalization:

A Method for Disaggregating Sector-level Digital Intensity to Firm-level

Tomasz Mucha[a]
Timo Seppälä[a, b]

Affiliations:

[a] Aalto University, Finland
[b] ETLA Economic Research, Finland

Original publication:

MethodsX 8 (2021) 101233.

Abstract

The digital transformation of firms plays an increasingly important role in the economy and society. However, limited access to data on firm-level digital intensity is an impediment to advancement of multiple research projects concerned with firm digitalization. To alleviate this challenge, this paper proposes a method for estimating firm-level digital intensity based on other more readily available firm-level data and reference data on digitalization, which is available on sector-level. The proposed method utilizes firm-level revenue breakdown by sector to estimate sector revenue-weighted digital intensity scores, which lead to classification of firms into low, medium and high digital intensity groups. The output from the proposed method can be directly used in research concerned with firm digitalization and investigating this multifaceted phenomenon. Results from the application of the proposed method to an illustrative sample of large US and non-US firms (2000 observations in total) indicate that firm-level digital intensity can be efficiently estimated for large samples using data commonly available to researchers.

The key differences between the proposed method and alternative methods are:

- Recognition of the fact that firms might participate in more than one sector or industry, which partially explains within-sector heterogeneity in firm-level digital intensity. We found that 67.8% of large US firms and 78.6% of large non-US firms were engaged in more than one industry.
- Use of reference sector-level digital intensity scores, which allows for rapid update, application across geographies and time, as well as parallel calculation of multiple digital intensity scores for each reference data. Furthermore, use of reference data enables supplementation of firm-level data on digitalization.
- Replicability of the method and reproducibility of the results through inclusion of the source code and availability of data through research and commercial databases.

Keywords

Digital transformation, Digital taxonomy, IT intensity, Data disaggregation

Method details

The digital transformation of firms plays an increasingly important role in the economy and society. Digitalization affects organizations from a variety of angles and levels [1]. Furthermore, this phenomenon impacts organizations across the full range of industries and sectors [2]. Hence, research on digitalization of firms and other phenomena related to it is of significant importance to the society. This observation is supported by increasing research interest in these topics across various disciplines [3]. Such research is enabled, but also potentially limited, by the extent of available methodological toolbox. Methods used in research on digitalization span a wide range, including both quantitative and qualitative methods [1]. These methods take a variety of data as inputs, such as case studies [1], aggregate measures of investment in information and communication technologies (ICT) stock [4], [5], purchases of intermediate ICT goods and services [6], robot use [7], [8], online sales [6], and occupational classification and task-based index of digital intensity [9]. However, due to the fact that "inherent difficulties exist in measuring the scope and pace of such a multifaceted phenomenon" [6, p. 5] as digitalization, access to suitable data might be an impediment to advancement of our understanding.

The present paper proposes a method, which alleviates the challenge of insufficient firm-level data by leveraging suitable results from past research on sector-level digitalization. The proposed method utilizes firm-level revenue breakdown by sector to estimate sector revenue-weighted digital intensity scores. These scores are derived from existing results of research on sector-level digitalization. The method output is a classification of firms into low, medium and high digital intensity groups.

The reminder of this paper is divided into three sections. We first discuss input data. After that we describe steps in the method and conclude with method validation. The paper is accompanied with supplementary material, which includes R code for implementation and validation of the method, as well as sample data used in the validation section.

Input data

The implementation of the proposed method relies on three categories of input data. First two are necessary, while the third one is used in special cases only. These categories are:

- Firm-level data on revenue per sector or industry.
- Reference sector-level digital intensity scores.
- Additionally, in case these two categories of data listed above rely on different industry classification systems, there is a need for a concordance table, which maps industry classification codes on a firm-level to those on a sector-level.

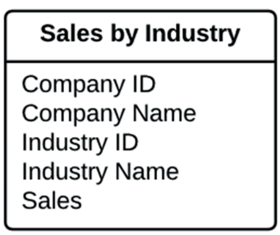
Firm-level data

Firm-level data is the data describing companies of interest. At a minimum, firm-level data must include firm-specific identifier, industry or sector code (thereafter, referred to as industry code, for brevity) and corresponding revenue or share of annual revenue. A single company might be active in either one or many industries. Additional information, such as firm name and industry name is useful to include to facilitate manual inspection of data processing steps, when in the development phase. Once the proposed method produces its outputs, these intermediary results will likely need to be combined with other data and subjected to analysis to address specific research questions.

It is important to recognize that the proposed method uses, for each company, revenue figures allocated to relevant industries as basis for calculating weights, which in turn are utilized to calculate revenue-weighted digital intensity score of each company. We motivate the use of revenue as the key determinant of industry participation with the following logic. Companies generating revenue from a given industry are likely to have characteristics similar to those of other companies in that industry. This is driven by similarity of the environmental conditions in which they operate, such as customer base, regulation, competition, technology context, etc. In summary, our argument for the use of revenue split by industry as a proxy for digital intensity score weights is based on the institutional isomorphism logic [10]. Thus, digital intensity of a company should, approximately, be the digital intensity of each industry where that company is active and proportional to the level of activity in these industries.

While the firm-level data can take a simple format, as presented in the Figure 1, it is common to encounter more complex input data and data issues. For example, there might be multiple industry codes grouped together and representing a single business segment of a company, which is accompanied by a single revenue figure. Another difficulty might be negative figures reported as eliminations resulting from inter-segment sales. Finally, industry classification systems have been periodically revised, thus it is possible to encounter industry codes from different revisions

Figure 1 An example of a simple data structure for firm-level data



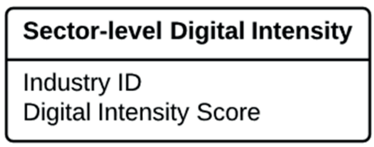
of an industry classification system listed in the same data set. We propose several sub-procedures for dealing with such data issues in the latter section of this paper. If other types of complexities are encounters, researchers must use common sense to process or convert the data to comply with the requirements of the latter steps in the procedure. Furthermore, any such judgement calls and additions to the procedure should be documented and reported.

Sector-level data

Sector-level data is the source data for digital intensity scores. Our method leverages previous research on the digital and IT intensity of industries, for example [6], [11]. Published results for digital intensity of industries serves as a reference data for the proposed method. This approach presents some limitations, which need to be recognized before application of the method. Scholars applying the proposed method in own research ought to assess the suitability of the sector-level reference data for the estimation of digital intensity on a firm-level for the specific sample of companies under investigation. Researchers need to evaluate the alignment between the two data sets considering multiple factors. First, the alignment in time frame needs to be assessed. Since digital intensity of sectors might be changing over time [6], it is important to evaluate whether the reference data is representative of the sample, given potential temporal changes in digital intensity. Next, there are differences in the level of sectoral digital intensity in different countries [6], thus overlap in geographic coverage needs to be considered. Firm size is another important aspect, as size is positively correlated with variables associated with digital technology adoption [4]. These variables include, but are not limited to, slack resources, access to finance, wealth, scale, and specialization [12], [13]. Another set of factors relate to market concentration and competitiveness, which can be assessed, for example, using Herfindahl-Hirschman index [14]. Market concentration and competitiveness are associated with adoption rates for high technology [12], [15], thus alignment between the reference data and the sample data needs to be assessed with this respect as well. Finally, the methodology used in the sector level analysis leading to the reference data should be evaluated for suitability with the research question at hand. Other factors potentially undermining the suitability of the reference data for use with the specific sample under investigation might need to be considered as well. Yet, given limited availability and difficulty with access to information needed for calculating digital intensity directly on a firm-level, use of a reference data on a sector-level presents a viable alternative. Furthermore, this approach enables researchers to estimate on a per-firm basis multiple digital intensity scores based on alternative reference data sources, as well as efficiently revise existing digital intensity scores when new reference data becomes available.

Sector-level digital intensity data takes the form of a simple look-up table with industry codes and their respective digital intensity scores, as presented in Figure 2. It is useful to retain industry names in the data to facilitate debugging of the procedure, while in the development phase. Potential complexities relate to the aggregation of multiple industries into ranges of industry codes. This might also be associated with some papers using industry codes on different levels in the taxonomy of an industry classification system. While simple aggregation of industries based on industry taxonomies are straightforward to handle, researchers developing sector-level digital intensity scores might also make discretionary decisions regarding aggregation into higher-level industries or sectors. In such cases it is important to evaluate and, potentially, disentangle earlier modifications to the industry classification taxonomy. Again, transparency and common sense need to be applied and choices documented.

Figure 2 The simplest possible format for a sector-level digital intensity score

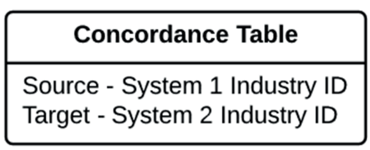


Concordance table

According to U.S. Census Bureau, concordance tables "provide detailed descriptions of the direct relationships between classification systems" [16]. These tables map industry codes from one industry classification system to another, as well as map industry codes within the same classification system for different revisions of that system. The data structure for concordance tables is presented in Figure 3.

In cases where the firm-level data or both firm- and sector-level data include industry codes from different industry classification systems or different revisions of

Figure 3 Data structure in a concordance table



the same system the use of concordance tables will be required in the application of the proposed method. Concordance tables are provided by national or international census or statistical offices and, therefore, tend to be a reliable, replicable and easily available. However, potential data issues might relate to translation of older industry classification systems into more recent ones. For example, U.S. Census Bureau does not provide direct concordance table between NAICS (North American Industry Classification System) 1997 to NAICS 2017. In the next section of this paper we discuss two approaches for dealing with such data issue.

Steps in the method for disaggregating sector-level digital intensity scores to firm-level

In this section, we first outline the steps involved in the implementation of the proposed method. Thereafter, we discuss each step and provide a commentary on how to deal with potential data issues.

The key steps in the implementation of the method are:

1. For each company retrieve data with or calculate revenue figure for each industry code.
2. In case firm-level and sector-level data uses different industry classification systems or different revisions of the same classification system, use concordance table(s) to convert firm-level industry codes to those at sector-level.
3. For each firm-level industry code match the corresponding digital intensity score using the sector-level data as a reference (look-up table).
4. For each company, calculate revenue-weighted digital intensity score.
5. For each company, classify the revenue-weighted digital intensity score into one of three digital intensity groups (low, mid or high).

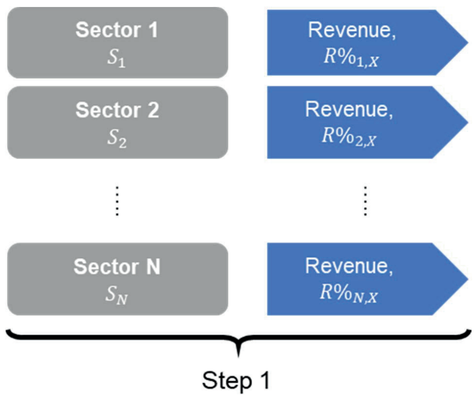
Step 1: Company revenue per industry code

Depending on the data source, the data might be readily available, or some data processing might be needed. Common data processing requirements include:

- Splitting business segment revenue to multiple industry codes
- Dealing with negative figures
- Dealing with missing revenue breakdown by business segment or industry

Since many companies provide information on their sales per business segment (typically, in annual reports in the notes to the financial statements) it is likely that

Figure 4 Step 1: For each company, retrieval or calculation of revenue stream broken down by sector



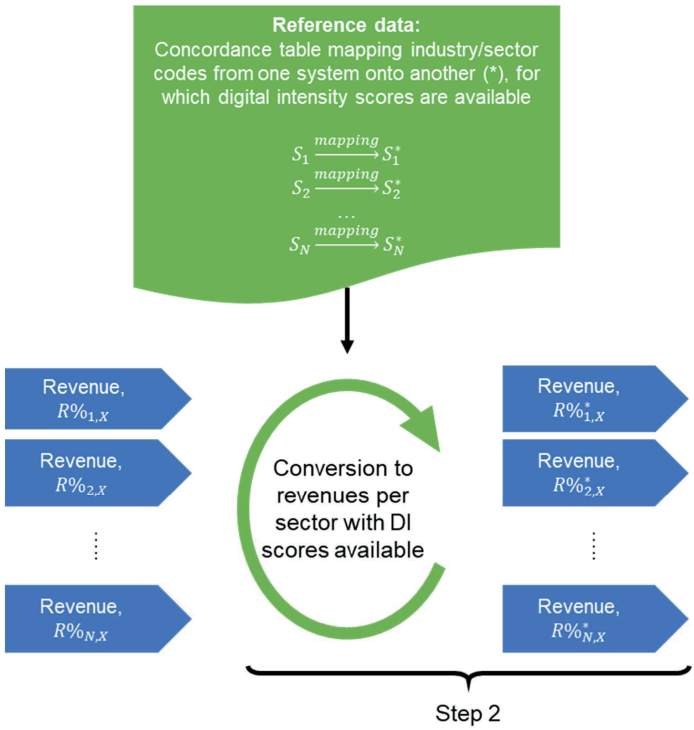
revenue data is recorded on a per business segment basis, rather than per industry code basis. Nevertheless, business segments can be matched with one or multiple industry codes. This can be done by researchers themselves or such information can be available in the financial databases. In either case, it is common to encounter multiple industry codes assigned to a single business segment. If this is the case, each business segment revenue should be evenly split between industry codes. The justification for such treatment is that typically there is not enough information to assign different weights to individual industry codes. Equal weights reflect equal treatment of all industry codes assigned to a single business segment.

Another data issue, which is sometimes encountered, is negative revenue reported as eliminations of inter-segment sales within a company. We recommend dropping the records with negative revenue, since revenue from each business segment excluding eliminations should sufficiently well reflect the level of company engagement in different industries.

Finally, some companies do not report revenue breakdown by segment and, thus, it might not be possible to get data on revenue per industry code for such companies. The proposed method requires at least one industry code, which is available on a company-level. Such industry code is generally available for any registered company in the form of primary industry code. In some cases, several industry codes might also be available on a firm-level. In either case, the treatment of these industry codes is equivalent to the base case situation, where revenue per business segment is available. The only difference is that instead of using revenue per business segment to allocate revenue per industry code, it is the total revenue of a company, which is used. Primary and secondary industry codes are available in multiple financial databases.

Step 2: Converting firm-level industry codes to sector-level codes using concordance tables

Figure 5 Step 2: Conversion of industry codes related to firm-level revenue streams into another industry classification system, for which sector-level digital intensity scores are available. This step is required only if the firm-level data and sector-level reference data are expressed using a different industry classification systems.



This step can be skipped, if both firm-level and sector-level industry codes are expressed using the same industry classification system and the same revision of that system. In other cases, there is a need to harmonize the industry codes on both levels. This is achieved with concordance tables. Once industry codes on firm- and sector-level are matched it is possible to map sector-level digital intensity scores to firm-level in the next step.

Concordance tables can be downloaded from websites of, for example, U.S. Census Bureau [16] or Eurostat [17]. The latter source refers to concordance tables as correspondence tables.

Since it is possible that some industry codes in a concordance table are mapped to more than one code in another system or revision of industry classification, our method requires adjustment of some of the company revenue per industry code figures, which were calculated in the previous step. In line with the logic regarding splitting segment revenue to industry codes, which was presented earlier, we propose the same treatment for cases where concordance tables map a single industry code to multiple codes in another industry classification system. This means that if the concordance table applied maps one industry code to many, our method evenly splits company revenue related to that industry code and allocates that value to the resulting industry codes in another classification system or revision.

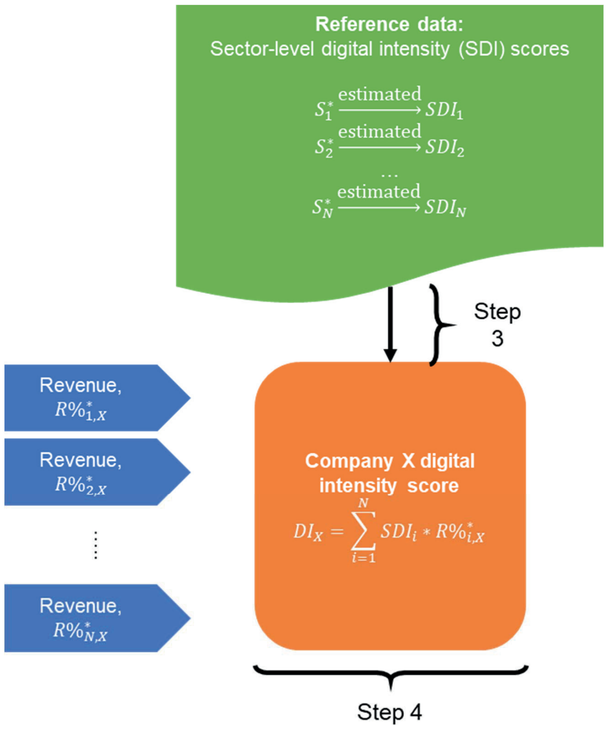
While the application of concordance tables, later revenue splitting and allocation of revenues to industry codes should be a straightforward procedure, there is one potential data issue, which reveals itself at this stage. In case the source industry codes are not all from the same revision of an industry classification system, it is possible that the concordance table applied does not map some of the source industry codes to any target industry code. This data issue can be resolved in two ways. Either (1) another concordance table can be used or (2) the same concordance table as previously can be used with both source and target industry codes escalated by one level in the industry classification taxonomy.

We recommend using the first approach, if concordance tables for other revisions of the source industry codes are available. This step can be repeated iteratively until all missing values are replaced with the corresponding target industry codes. Alternatively, and preferably after applying multiple concordance tables, the remaining missing values can be replaced with target industry codes by using the second approach proposed.

In the second approach, the original concordance table is modified by dropping the last digit in the industry codes (both source and target). Also, the firm-level industry codes need to be generalized in the same way. At this point it is important to recognize that dropping the last digit in the industry codes might result in some firm-level records appearing as duplicates. These duplicates appear due to some firm-level records differing between each other only with the last digit of the industry code. If such duplicates appear, they should be merged by summing the revenue figure for all records that are duplicates of each other and removing all, but one. Once this is completed the more generalized concordance table can be reapplied to the more generalized firm-level industry codes. This approach can be iteratively applied until all missing values are replaced with target industry codes.

Step 3: Mapping firm-level industry codes to sector-level digital intensity scores

Figure 6 Step 3: Matching of revenue streams and their corresponding industry codes with sector-level digital intensity scores, which come from reference data. Step 4: Calculation of revenue-weighted digital intensity score for each company.



Given that both firm- and sector-level industry codes are expressed using the same industry classification system and its revision, mapping digital intensity scores, which are at sector-level, to industry codes on a firm-level is a matter of using a simple look-up table logic. There should be no data issues present at this stage. However, it is important to validate that there are no missing values, which could result from incomplete industry code coverage of the sector-level digital intensity scores.

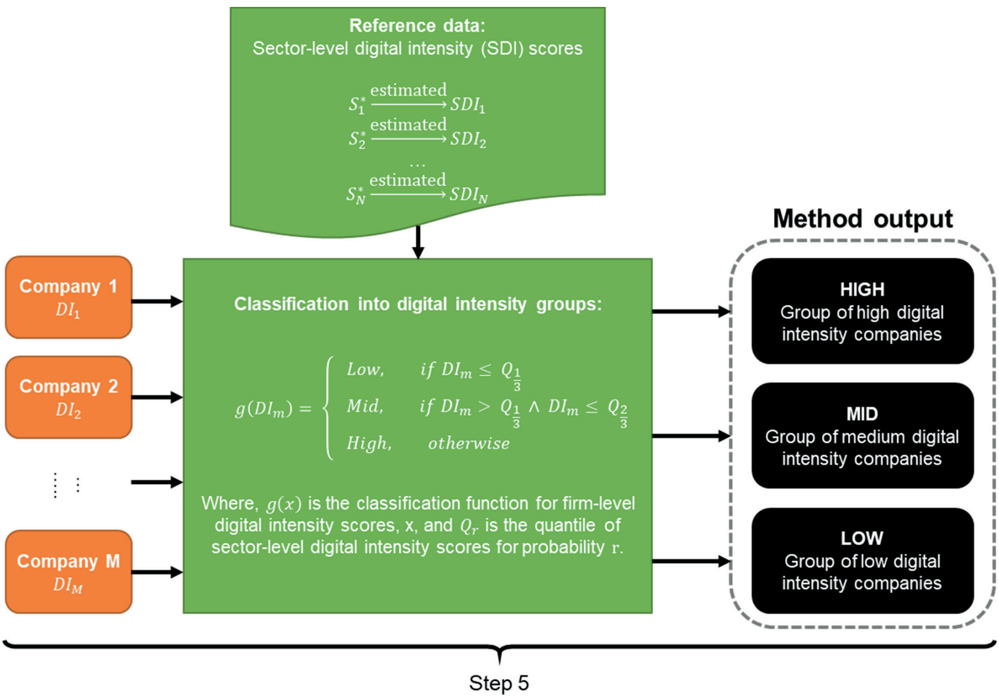
Step 4: Firm-level revenue-weighted digital intensity score

Once sector digital intensity scores, SDI_i, are available at firm-level for each industry code, i, the final digital intensity score, DI_X, for company X is calculated as a weighted average of sector digital intensity scores SDI_i (Equation 1.), where weights, $R\%^*_{i,X}$, are expressed as share of company X revenue coming from industry i. Star in $R\%^*_{i,X}$ denotes that the revenue share is for the industry code i, which is expressed in the same industry classification system and revision of that system as that of the sector digital intensity score SDI_i.

$$DI_X = \sum_{i=1}^{N} SDI_i * R\%^*_{i,X} \tag{1}$$

Step 5: Classification of digital intensity scores into three groups

Figure 7 Step 5: Classification of firms into digital intensity groups based on firm-level digital intensity scores and using cut-off points (quantiles with probabilities 1/3 and 2/3) based on reference sector-level digital intensity scores.



The final step is classification of firm-level revenue-weighted digital intensity scores into low, medium, and high digital intensity groups. This step is important because of two reasons. First, since the proposed method disaggregates sector level generalizations to firm level, it is an imperative to recognize that the assigned firm-level digital intensity scores cannot be considered as precise figures. Calvino and colleagues [6] report high level of within-sector heterogeneity for many of the digital intensity indicators they consider. Furthermore, they highlight that there can be many alternative ways to aggregate digital intensity indicators into a "global" indicator. This methodological ambiguity reflects the complexity of the underlying phenomenon. Given that digitalization itself is multifaceted, complex, and evolving we do not expect that a single method can fully capture that phenomenon. Second, the proposed method is intended for use with both ordinal and ratio sector-level digital intensity score scales. The lower information content in ordinal scales creates the requirement for simplification of the final method outputs. Overall, given the two reasons discussed above, we consider that the proposed method strikes the right balance between providing useful granularity and acceptable risk of misclassifying companies.

$$g(DI_m) = \begin{cases} Low, & if\ DI_m \leq Q_{\frac{1}{3}} \\ Mid, & if\ DI_m > Q_{\frac{1}{3}} \wedge DI_m \leq Q_{\frac{2}{3}} \\ High, & otherwise \end{cases} \tag{2}$$

The classification of firm-level digital intensity scores, DI_m, into groups is carried out using a classification function $g(x)$, where Qr is the quantile of reference sector-level digital intensity scores for probability r. The cut-off values between the groups are calculated from the reference data rather than from the firm-level digital intensity scores calculated in Step 4, because there is no guarantee that the sample of companies under analysis is representative of the whole economy. Reference data, on the other hand, is more likely to meet this requirement.

Method validation

Firm-level data

We apply the proposed method to estimate digital intensity scores for two samples of companies. Both selected samples include 1000 largest companies (based on market capitalization), as of 31st August 2020 and based on country of headquarters:

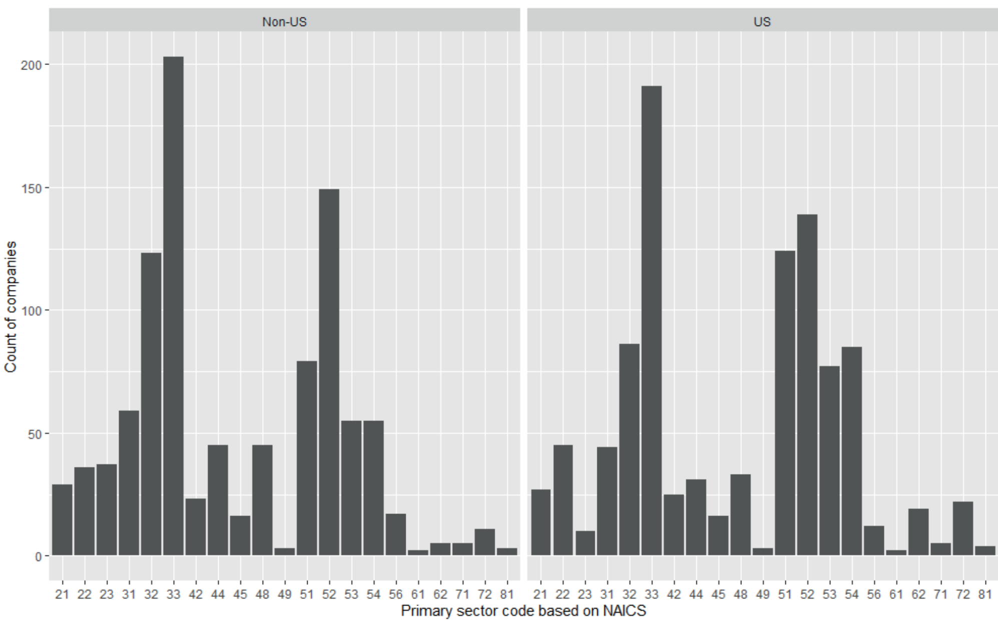
- US Sample: Companies headquartered in the U.S.
- Non-US Sample: Companies headquartered in Australia, Austria, Denmark, Finland, France, Italy, Japan, the Netherlands, Norway, Sweden, and the United Kingdom.

We retrieved the firm-level data from Thomson Reuters Eikon database. For each sample, the distribution of company count by two-digit NAICS code is presented in Figure 9. We used Eikon Screener App to find unique identifiers (RICs) of publicly listed companies based on respective country of headquarters and market capitalization denominated in USD. Furthermore, we excluded all ETFs (Exchange Traded Funds) and closed-end funds from the sample. We then used Thomson Reuters MS Excel Add-In to retrieve for each company the following items:

- company name
- primary industry code (North American Industry Classification, NAICS)
- primary industry name
- segment code (NAICS)
- segment name
- business total revenue by segment

The samples of companies used in this section were selected for illustrative purposes only. The use of the method is not restricted to countries included in this analysis nor to large companies only. As discussed in the Input data section of this paper, it is the choice of the reference data that determines suitability of the proposed method for the specific sample of companies under investigation. We discuss reference data used in this analysis in the following section.

Figure 8 Count of companies by sector (based on first two digits of primary NAICS code)



Reference data

Analysis of our samples required two types of reference data, which were concordance tables and sector-level digital intensity scores. Since industry codes available in the firm-level data (NAICS codes) and sector-level data (ISIC codes) were expressed using different classification systems, we needed to employ concordance tables to translate between them. We relied on concordance tables mapping NAICS codes to ISIC codes available from U.S. Census Bureau [16]. Furthermore, since some NAICS codes were expressed using revisions of NAICS classification other than the latest, 2017 revision, in some cases we needed to map these older NAICS to more recent revisions of NAICS. This mapping was also done using concordance tables available from the same source. Sector-level digital intensity scores are discussed in more detail in the remainder of this section.

The 12 countries, which are covered by the sample, were selected, because they are included in the OECD taxonomy of digital intensive industries [6], which is the source of our reference data covering sector-level digital intensity scores. We consider that this reference data is a good example of input that is suitable for the proposed method. In case of OECD taxonomy, digitalization is considered through multiple indicators, thus capturing numerous facets of this complex phenomenon. Other alternative sector-level digital intensity scores, such as those calculated by Brynjolfsson and colleagues [11], could be used as well, although alignment of the selected samples and the reference data would not be as good due to differences in geographic coverage. Users of the proposed method must decide which reference data for sector-level digital intensity is suitable for their research question and design.

Despite the fact that Calvino and colleagues [6] do not report sector-level digital intensity scores directly in their paper, we can replicate their ultimate "global" taxonomy results for all, but one sector, thus achieve 97.22% agreement between our results. Based on our calculation of "global" sector-level digital intensity scores "Transport equipment" sector falls into one digital intensity group lower than what is presented in the results of Calvino and colleagues [6]. We attribute the difference in our replication results to the fact that our classification of sectors into groups of "global" indicator for digital intensity relies on indicator-level digital intensity scores aggregated across countries and years (this data is openly available from OECD via a StatLink dx.doi.org/10.1787/888933617434). Thus, variability on country- or year-level could lead to different classification of "Transport equipment" sector. Nevertheless, we consider that the high degree of alignment between our results is sufficient to rely on our estimation of sector-level digital intensity scores in the reminder of the analysis. The sector-level digital intensity scores used in this analysis are presented in Table 1 and are also available for download from the supplementary materials available with this article.

These sector-level digital intensity scores are used in the analysis as a reference look-up table for assigning digital intensity scores to company-level streams of reve-

Table 1 Reference data for sector-level digital intensity scores

Sector	ISIC code (rev. 4)	Digital Intensity Score*
Agriculture, forestry, fishing	01–03	0,0463
Mining and quarrying	05–09	0,2361
Food products, beverages and tobacco	10–12	0,3254
Textiles, wearing apparel, leather	13–15	0,4246
Wood and paper products, and printing	16–18	0,4563
Coke and refined petroleum products	19	0,3532
Chemicals and chemical products	20	0,4087
Pharmaceutical products	21	0,3651
Rubber and plastics products	22–23	0,4365
Basic metals and fabricated metal products	24–25	0,3690
Computer, electronic and optical products	26	0,5648
Electrical equipment	27	0,5185
Machinery and equipment n.e.c.	28	0,5324
Transport equipment	29–30	0,6157
Furniture; other manufacturing; repairs of computers	31–33	0,5754
Electricity, gas, steam and air cond.	35	0,3016
Water supply; sewerage, waste management	36–39	0,3016
Construction	41–43	0,2698
Wholesale and retail trade, repair	45–47	0,5926
Transportation and storage	49–53	0,3194
Accommodation and food service activities	55–56	0,2870
Publishing, audiovisual and broadcasting	58–60	0,6157
Telecommunications	61	0,8796
IT and other information services	62–63	0,8241
Finance and insurance	64–66	0,8222
Real estate	68	0,0741
Legal and accounting activities, etc.	69–71	0,6620
Scientific research and development	72	0,6204
Advertising and market research; other business services	73–75	0,6806
Administrative and support service activities	77–82	0,6528
Public administration and defence	84	0,5333
Education	85	0,3944
Human health activities	86	0,4333
Residential care and social work activities	87–88	0,4111
Arts, entertainment and recreation	90–93	0,4889
Other service activities	94–96	0,6167

* These scores were estimated following the methodology developed by Calvino and colleagues [6] and using data available from OECD via a StatLink dx.doi.org/10.1787/888933617434. The scores themselves do not have direct interpretation other than providing ranking of sectors in terms of their digital intensity.

nue coming from activity in different sectors. Once revenue-weighted digital intensity scores are calculated for each company, we use again the reference look-up table to compare these scores against cut-off points between low, medium and high digital intensity sectors. These cut-off points are quantiles in the reference look-up table digital intensity scores corresponding to 1/3 and 2/3 probabilities. Thus, given our reference data, firms with revenue-weighted digital intensity score below 0.386 are classified as low digital intensity, those with scores above 0.568 are classified as high digital intensity, and those in between are medium digital intensity.

Efficiency of the method

Using the input data and following the proposed method (steps 1–5) yields a classification of firm-level digital intensity into three groups as presented in Figure 9.

In the absence of any other information, the input data was enough to estimate digital intensity for the sample companies on a firm-level, thus demonstrating the efficiency of the proposed method, given suitable sector-level reference data is available. Relatively low data requirement and accessibility of the required data make the proposed method practically feasible for use. Such data efficiency is the primary advantage of the proposed method, which despite the lack of more detailed data on company digitalization can be used in a wide range of research work.

Figure 9 Visual representation of method output for the two data samples

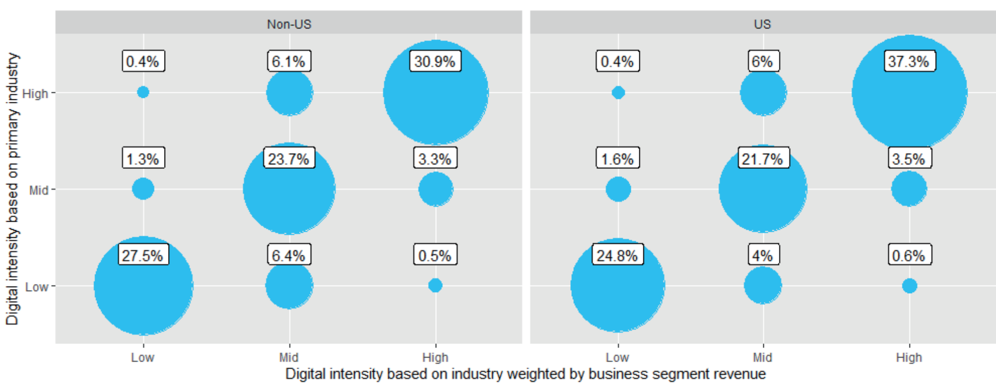
Another aspect of the proposed method is procedural clarity, which leads to higher replicability and comparability in the studies investigating or utilizing firm-level digitalization measures. Not only should the description of the proposed method provided in this paper be used to inform researchers regarding the method steps, but also R code included in the supplementary material should provide means for higher replicability.

Finally, given the automation of data processing using R script and separation of the method inputs into firm-level data and reference data, this method provides means for research updatability. Once new firm-level data or reference data on sector-level digital intensity becomes available, the requirement for resources needed to recalculate and update results is low.

Comparison of firm digital intensity based on primary industry only and segment level industries

As we noted in the description of Step 1 of the proposed method, primary industry codes can be used to supplement the firm-level data in cases where revenue breakdown by business segment is not available for some companies. However, it is important to point out that there is a potential trade-off related to inclusion of companies with lacking data on segment revenue. While it is likely that researchers applying the proposed method will not have full coverage of firm-level business segment revenue data for their samples, we would recommend using the proposed method only in cases where majority of the sample has such data available. To demonstrate the difference in the results, which are based on data with full access to business seg-

Figure 10 Comparison of method results with and without firm-level business segment revenue data



ment revenue and data with primary industry codes only, we provide comparative results in this section.

We used the input data consisting of the same two samples as in the previous section as the starting point for this analysis. After excluding companies, which did not have revenue breakdown by business segment, were left with 678 and 786 observations for US and Non-US samples, respectively.

Using these restricted samples, we recalculated the results of the proposed method. We refer to these results as “digital intensity based on industry weighted by business segment revenue”. Thereafter, we removed business segment revenue information from the restricted samples and recalculated the results. Since this second application of the proposed method could not use business segment revenue as weights to calculate firm-level digital intensity, only information regarding primary industry of each company was used. We refer to these results as “digital intensity based on primary industry”. Comparison of the results from both runs is presented in Figure 10.

There is an overall alignment between the results from each calculation run, as presented in Table 2. Cohen’s kappa for both samples is relatively high, thus indicating agreement between the two approaches. However, this result is expected, as the null hypothesis for Cohen’s kappa is random grouping of the observations. In our case, we are more interested to detect if there is difference between the two approaches in terms of groupings. While simple percentage agreement is above 80% for both samples, the permutation test rejects, at 5% significance level, the hypothesis that the agreement is 100%. Thus, we conclude that there is higher information content in the approach relying on business segment revenue figures and the resulting revenue-weighted digital intensity scores. Our recommendation is to use to the extent possible firm-level data with information on revenue per industry or business segment. In our view, this is a superior approach to one relying solely on primary industry codes.

Table 2 Agreement in classification of companies into digital intensity groups between results with and without firm-level business segment revenue data

Sample	Non-US	US
Observation count	786	678
Cohen’s kappa	0.731	0.752
Simple percentage agreement	82.1%	83.8%
	(79.3%, 84.7%)*	(81.0%, 86.4%)*

* Values in parenthesis show estimated confidence interval for a = 5% using permutation test with 5000 bootstraps.

Conclusion and limitations

Overall, the proposed method exhibits the key intended property, which is efficient estimation of firm-level digital intensity, while utilizing data that is readily available for large samples of companies. By leveraging information on the level of business activity of companies in different industries and sectors the proposed method allows scholars to tap into results from previous research on digital intensity of sectors. The results from validation of the method against two samples of companies with 1000 observations each reveal that classification of firms into low, medium and high digital intensity groups is significantly different from alternative classification, where only information on firm primary industry is used. Thus, we conclude that the proposed method using revenue-weighted digital intensity scores produces superior estimates of firm digital intensity.

Since the proposed method relies on sector-level reference data on digitalization, its results can be only as good as the quality of the reference data. While this presents a limitation, it provides also a benefit in the form of updatability of the research results. Simply swapping the reference data to a different or newer version, with no further alternations in the estimation procedure, generates potentially more appropriate or more up-to-date results. This means that the proposed method is flexible in the sense that researchers can choose reference data to match the geography, timeframe and other parameters of their firm-level data. Furthermore, even if firm-level data on digitalization is available to some extent, for example covering only certain aspects of digitalization, the proposed method can be used to augment or supplement the data, thus potentially providing better operationalization of firm digitalization.

Finally, the proposed method is intended to increase transparency and replicability of research on digitalization. The supplementary material included with this paper comprises of not only input data used in the method validation section, but also source code (in R language), which allows for exact reproduction of the results. Thanks to the source code and relative availability of input data, which is suitable for the proposed method, large samples of companies can be classified into digital intensity groups in a manner, which is transparent to the research community.

The proposed method can also be further developed to incorporate other measures of firm engagement in different sectors. For example, apart from relying on revenue as an indicator of sector engagement, sourcing relationships could also provide useful input to the method. Analysis of sourcing relationships allow for derivation of value-add distribution across supply chain [18]–[20]we perform grass-roots investigative work to uncover the geography of the value added for a Nokia N95 smartphone circa 2007. The phone was assembled in Finland and China. When the device was assembled and sold in Europe, the value-added share of Europe (EU-27 and thus could provide an up-stream perspective on digitalization.

Supplementary material:

- R code
- Report generated from R, where all the calculation steps and outputs are visible
- Sample data:
  - Revenue breakdown by business segment for 2000 sample companies, primary and business segment industry codes
  - Reference data with sector level digital intensity scores and taxonomy
  - Concordance tables mapping different industry classification systems and evisions to each other
  - Sector-level digital intensity scores calculated based on OECD data available from dx.doi.org/10.1787/888933617434

References

[1] E. Henriette, M. Feki, and I. Boughzala, "The Shape of Digital Transformation: A Systematic Literature Review," 2015.

[2] J. Reis, M. Amorim, N. Melão, and P. Matos, "Digital transformation: a literature review and guidelines for future research," in *World conference on information systems and technologies*, 2018, pp. 411–421.

[3] R. Morakanyane, A. A. Grace, and P. O'Reilly, "Conceptualizing Digital Transformation in Business Organizations: A Systematic Review of Literature.," in *Bled eConference*, 2017, p. 21.

[4] D. Pilat and F. Lee, "Productivity Growth in ICT-producing and ICT-using Industries: A Source of Growth Differentials in the OECD?," 2001.

[5] A. Colecchia and P. Schreyer, "ICT investment and economic growth in the 1990s: is the United States a unique case?: a comparative study of nine OECD countries," *Rev. Econ. Dyn.*, vol. 5, no. 2, pp. 408–442, 2002.

[6] F. Calvino, C. Criscuolo, L. Marcolin, and M. Squicciarini, "A taxonomy of digital intensive sectors," 2018, [Online]. Available: https://www.oecd-ilibrary.org/docserver/f404736a-en.pdf.

[7] W. Dauth, S. Findeisen, J. Südekum, and N. Woessner, "German robots-the impact of industrial robots on workers," 2017.

[8] D. Acemoğlu and P. Restrepo, "The race between machines and humans: Implications for growth, factor shares and jobs," *Retrieved*, vol. 6, p. 2019, 2016.

[9] G. Gallipoli and C. Makridis, "Sectoral Digital Intensity and GDP Growth After a Large Employment Shock: A Simple Extrapolation Exercise," *SSRN Electron. J.*, 2020, doi: 10.2139/ssrn.3660598.

[10] P. J. DiMaggio and W. W. Powell, "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields," *Am. Sociol. Rev.*, pp. 147–160, 1983.

[11] E. Brynjolfsson, A. McAfee, M. Sorell, and F. Zhu, "Scale without mass: business process replication and industry dynamics," *Harv. Bus. Sch. Technol. Oper. Mgt Unit Res. Pap.*, no. 07–016, 2008, [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=980568&mod=article_inline.

[12] R. G. Fichman, "The diffusion and assimilation of information technology innovations," *Fram. Domains IT Manag. Proj. Future Past*, vol. 105127, pp. 105–128, 2000.

[13] T. Greenhalgh, G. Robert, P. Bate, F. Macfarlane, and O. Kyriakidou, *Diffusion of innovations in health service organisations: a systematic literature review*. John Wiley & Sons, 2008.

[14] S. A. Rhoades, "The herfindahl-hirschman index," *Fed Res Bull*, vol. 79, p. 188, 1993.

[15] H. Gatignon and T. S. Robertson, “Technology diffusion: an empirical test of competitive effects,” *J. Mark.*, vol. 53, no. 1, pp. 35–49, 1989.

[16] C. D. B. US Census Bureau, “North American Industry Classification System (NAICS), Concordances - US Census Bureau.” https://www.census.gov/eos/www/naics/concordances/concordances.html (accessed Jun. 05, 2020).

[17] “Europa - RAMON - Correspondence Tables List.” https://ec.europa.eu/eurostat/ramon/relations/index.cfm?TargetUrl=LST_REL&StrLanguageCode=EN&IntCurrentPage=11 (accessed Jun. 05, 2020).

[18] J. Ali-Yrkkö, P. Rouvinen, T. Seppälä, and P. Ylä-Anttila, “Who Captures Value in Global Supply Chains? Case Nokia N95 Smartphone,” *J. Ind. Compet. Trade*, vol. 11, no. 3, pp. 263–278, Sep. 2011, doi: 10.1007/s10842-011-0107-4.

[19] T. Seppälä, M. Kenney, and J. Ali-Yrkkö, “Global supply chains and transfer pricing: Insights from a case study,” *Supply Chain Manag. Int. J.*, vol. 19, no. 4, pp. 445–454, Jan. 2014, doi: 10.1108/SCM-01-2014-0049.

[20] M. M. Larsen, T. Seppälä, and J. Ali-Yrkkö, “The changing geography and ownership of value creation: evidence from mobile telecommunications,” *Ind. Innov.*, vol. 25, no. 7, pp. 675–698, Aug. 2018, doi: 10.1080/13662716.2017.1329086.

ARTICLE 2

AI Diffusion Monitoring among S&P500 Companies:

Empirical Results and Methodological Advancements

Tomasz Mucha[a]
Timo Seppälä[a, b]

Affiliations:

[a] Aalto University, Finland
[b] ETLA Economic Research, Finland

Original publication:

Mucha Tomasz and Seppala Timo, "AI Diffusion Monitoring among S&P500 Companies: Empirical Results and Methodological Advancements" (2022). DIGIT 2022 Proceedings. 13.

Abstract

With the increasing pace of digital technology innovation and commercialization, monitoring commercial diffusion of technologies becomes more important for organizations. Technology monitoring is fundamental to R&D planning, technology management, and strategic decision-making. Despite its importance, monitoring the diffusion of technologies at the commercial lifecycle stage relies on crude methods, such as “snapshot-in-time” surveys and keyword counts. These approaches are in stark contrast to novel and rapidly advancing methods for monitoring technologies at the precommercial lifecycle stages, such as fundamental scientific research and applied R&D. We address this imbalance by proposing a specialized method for monitoring the commercial diffusion of technology. The method recognizes phases in technology adoption by organizations and captures the temporal progression of the diffusion process. One of the central elements of the proposed method is the classification of text, which relies on qualitative content coding. Our approach to coding leverages the insights from innovation diffusion research and is sensitized specifically to detect phases in technology adoption by organizations. The approach is illustrated with the case of artificial intelligence (AI) diffusion among S&P 500 companies during the 2004–2019 period. Our first contribution is a new method for monitoring the commercial diffusion of technologies. It provides transparent, replicable, updatable, and granular results, which can complement survey-based technology monitoring. The second contribution is empirical evaluation of AI diffusion in the context of leading firms in North America.

Keywords

Technology diffusion, Technology adoption, Technology strategy, Artificial intelligence (AI), Machine learning (ML)

Introduction

Technology adoption is a fundamental driver of productivity and competitiveness for firms and nations (Brynjolfsson et al. 2018; Hall 2004). Hence, technology monitoring underlies the generation of strategic foresight regarding changes impacting businesses, economies, and societies (Roper et al. 2011, secs. 1 and 4.2). Therefore, monitoring technologies throughout their lifecycles is highly relevant to both research and practice. The method with the longest track record and commonly used today is survey-based research (Roper et al. 2011, pp. 100–103). Survey-based research is particularly prevalent in studies concerned with technologies entering commercialization and later stages in the technology lifecycle. Scholars and practitioners tasked with technology monitoring rely on surveys (for example, see: Balakrishnan et al. 2020; Magoulas and Swoyer 2020; Montagnier and Ek 2021; Oliveira et al. 2019; Zolas et al. 2020). Another group of technology monitoring methods, sometimes referred to as "tech mining" (Porter and Cunningham 2004), emerged from the content analysis (Roper et al. 2011, p. 106) and is currently under active development (Cunningham and Kwakkel 2016), particularly for monitoring precommercial-stage technologies. These novel approaches provide an increasing range of insights and inform R&D and technology planning related to precommercial-stage technologies.

Despite the contribution of these methods, limitations prevail in monitoring the commercial diffusion of technologies. First, since significant hurdles separate technological inventions and applied R&D from commercialization (Roper et al. 2011, p. 8), methods focused on early stages of the technology lifecycle, such as patent analysis, are not sufficient to understand the subsequent commercial diffusion of technology. Second, methods focused on later stages in the technology lifecycle also face limitations (Rogers 1983, p. 117). Thus, the development of monitoring methods suitable for commercial-stage technologies, which are longitudinal and recognize the complexity of the technology adoption process by organizations, has been missing. Therefore, we propose a method specifically designed to recognize phases in technology adoption by organizations and capture the diffusion process over time. The proposed method leverages the qualitative content analysis approach. Our approach to coding is sensitized to studying the organizational adoption of technologies. It builds on insights from innovation diffusion research concerned with the process of technology adoption within organizations (Cooper and Zmud 1990; Greenhalgh et al. 2008, sec. 5.3; Meyer and Goes 1988; Rogers 2010). The method is illustrated with the case of artificial intelligence (AI) diffusion among S&P 500 companies during the January 2004–May 2019 period. AI is a "frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems" (Berente et al. 2021, p. 1435). Top executives internationally recognize AI as having the potential to significantly impact the strategic position of their organizations and the competitive dynamics of industries (Ransbotham et al. 2020). Therefore, business leaders, scholars, and policy-makers are keen to monitor the commercial diffusion of AI.

This study brings several contributions. First, our method addresses the issue of technology monitoring for technologies in the latter part of their lifecycle, namely, those entering into commercialization or later stages. Second, the proposed method is versatile in terms of its applicability to a wide range of technologies. This versatility stems from its reliance on qualitative content analysis, which is not limited to any specific type of text or document, and its sensitization to broadly defined phases of technology adoption. Next, practitioners can readily adopt the proposed method into use and complement their existing technology monitoring approaches. Their projects will benefit from transparent, replicable, updatable, and granular results generated by our method. Thus, the proposed method presents a valuable addition to a survey-based approach to monitoring the commercial diffusion of technology. Finally, given that the proposed method follows a structured procedure for content coding, it may serve in the future as a foundation for an automated technology monitoring algorithm.

Theoretical background

Our approach draws on the existing research on technology monitoring and innovation diffusion. Therefore, in this section, we provide a brief overview of the relevant theory and methods from these two partially overlapping streams of literature. We separately identify the development of methods for monitoring the precommercial and commercial diffusion of technologies in both streams of literature. Technology monitoring is the process of observing and keeping up with developments in a specific technology (Roper et al. 2011, p. 72). It is widely used and provides essential inputs for both business and policy decision-makers and, thus, contributes to R&D management, technology management, and corporate and national strategies (Burgelman et al. 2004, pp. 8–9; Chen and Small 1994; Porter and Detampel 1995; Teichert and Mittermayer 2002). In this paper, we limit the scope of technology monitoring to include past developments.

Monitoring precommercial diffusion of technology

Companies cannot use precommercial-stage technologies in their daily operations but might engage with these technologies through, for example, R&D work. Nevertheless, understanding the development paths for precommercial technologies might be strategically important (Teichert and Mittermayer 2002). Since technological progress at the precommercial stage manifests itself, at least partially, in scientific publications and patents (Porter and Cunningham 2004, p. 7), technology monitoring primarily leverages these documents (Martino 2003; Roper et al. 2011, pp. 81–82). It is often referred to as "tech mining" (Porter and Cunningham 2004, sec. 2.3; Roper et al. 2011, sec. 5.2). Porter and Cunningham (2004, p. 19) define tech mining as "the application of text mining tools to science and technology information, informed by understanding of technological innovation processes." Particularly in the area of patent analysis, there have been many recent advances, such as analysis of innovation topics (Choi et al. 2018); identification of interindustry technologies (Fredström et al. 2021); and screening ideas in the early stages of technology development (Hong et al. 2021).

Despite these advances, monitoring the precommercial diffusion of technology is insufficient to understand the subsequent commercial diffusion. First, not all inventions "find a viable commercial application" (Grant 2016, p. 243). Next, there is a significant time lag between making an invention and its commercialization (Roper et al. 2011, sec. 1.2), which results from an innovation needing to overcome, in many cases, significant difficulties before the adopters take it into use (Rogers 2010, p. 1). Consequently, the methods for monitoring commercial diffusion of technology present a distinct area of research and practice.

Monitoring commercial diffusion of technology

Monitoring the commercial diffusion of technology develops an understanding of the extent to which the target population of adopters has taken a focal technology into use. It presents a unique set of challenges. Unlike in the case of precommercial diffusion, there are no commonly used and standardized publications to measure progression. Instead, technologies diffusing in a target market spread through various channels, such as industry conferences, press, word-of-mouth, business intelligence, and many more (Rogers 2010, pp. 18–20). Consequently, many types of actors engage in monitoring the commercial diffusion of various technologies. They include national statistical offices, not-for-profit organizations, and other service providers, such as market research firms and consulting companies.

There are two main categories of methods used in monitoring the commercial diffusion of technology: (1) survey research and (2) analysis of various types of content. We provide a brief background on the two categories and discuss their limitations. The use of surveys to collect data for research and analysis of commercial diffusion of technology has been and continues to be very prominent. A seminal study of hybrid corn diffusion in Iowa (Ryan and Gross 1943), which relied on interview-based surveys, formed the foundation of the diffusion research paradigm in the 1940s (Valente and Rogers 1995). Given the successful expansion of diffusion research in the following decades beyond the discipline of rural sociology (Rogers

2010), the previously established methodological approach continued to thrive and evolve. Recent studies investigating the diffusion of digital technologies continue to rely on surveys as a source of data (Oliveira et al. 2014, 2019). National statistical offices also use this approach to gauge the commercial diffusion of technology. A recent publication of U.S. enterprise technology adoption by the U.S. Census Bureau is a good illustration (Zolas et al. 2020). Additionally, major consulting companies and other organizations publishing insights on technology diffusion continue to rely on surveys, (for example, see: Balakrishnan et al. 2020; Magoulas and Swoyer 2020; Ransbotham et al. 2020). Despite this long lineage, survey research faces many limitations for providing insights into technology monitoring. Rogers (2010, pp. 126–130) highlights some of the criticism of survey-based methods. One of the limitations of surveys, which he points out, is providing a "snapshot-in-time" perspective rather than a "moving pictures" perspective. This low temporal granularity is a drawback, especially for rapidly advancing and diffusing technologies. Even remedying this by running surveys at multiple points in time introduces new challenges – distortion of the perception of innovation by the respondents (Rogers 1983, p. 117) and aggravation of nonresponse bias (Roper et al. 2011, p. 103). In addition, survey research in technology monitoring can suffer from long time lags, problems with definitions of technical terminology, and in the case of commercially run studies, limited transparency regarding specific methods and sampling (Montagnier and Ek 2021). Consequently, survey-based methods alone are not sufficient for monitoring the commercial diffusion of technology.

The second category of methods used in monitoring commercial diffusion of technology originates from content analysis. A study of the diffusion of multidivisional administrative structure among large industrial firms (Teece 1980) relied on qualitative content analysis. Teece analyzed, among others, annual reports, 10-K forms filed with the Securities and Exchange Commission, prospectuses, business periodical articles, recruiting literature, and publicly available texts of speeches by corporate officials. The resulting classification of organizational forms did not allow for multiple phases in innovation adoption but rather was binary in nature (Armour and Teece 1978). Similarly, a more recent analysis (Daniel Zhang et al. 2021, p. 106) disregarded phases in technology adoption by employing counts of technology-related keywords in executive presentations as an indicator of technology diffusion. Two other studies (Mikova and Sokolova 2019; Segev et al. 2015) analyzing the commercial diffusion of technology and employing content analysis also faced limitations, which resulted from the lack of control over the sample of companies included in the data analysis. Overall, we conclude that the current state of methods for monitoring the commercial diffusion of technology has been insufficient and stagnant. It is possible to address this gap by drawing on insights from innovation diffusion research on the process of technology adoption in firms (Greenhalgh et al. 2008, sec. 5.3; Rogers 2010, pp. 126–130).

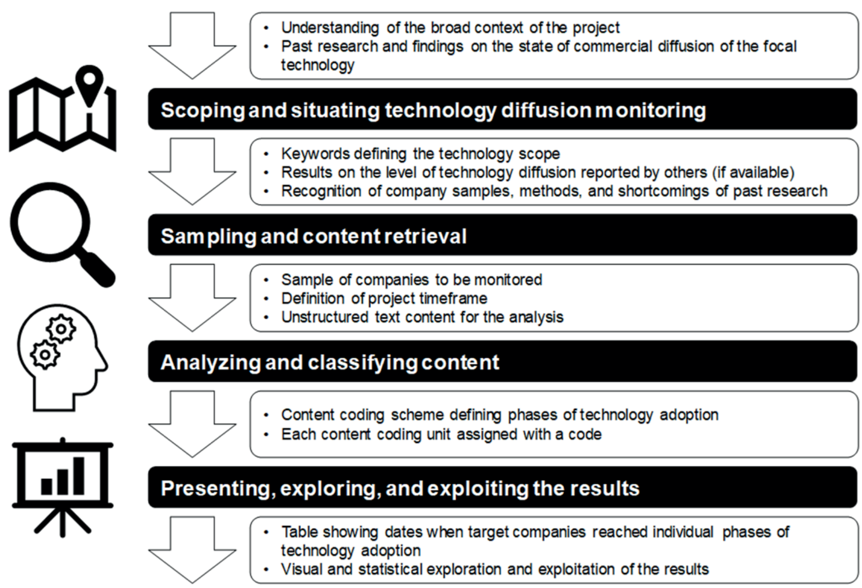
Proposed method

Our proposed approach to monitoring commercial diffusion of technology consists of four steps (Figure 1): (1) scoping and situating technology diffusion monitoring project, (2) sampling and content retrieval, (3) analyzing and classifying content, and (4) presenting, exploring, and exploiting the results. We describe these steps in greater detail in the following subsections.

Step 1: Scoping and situating technology diffusion monitoring project

The first task of researchers employing the proposed method is to define the scope of the monitoring project by identifying categories, names, or keywords representing the target technology. These keywords guide the content search and retrieval (in step 2). The terminology for describing and referring to (early) commercial-stage technologies is either established or emerging (Santos and Eisenhardt 2009). The trade-off between specificity and breadth of these keywords drives the scope of the monitoring project. For example, some keywords might represent a broader technological trajectory or frontier, such as "solar energy." Others might encompass only a narrower set of technologies, such as "tower concentrating solar plants." Furthermore,

Figure 1 Overview of steps in the proposed method



the level of project scoping difficulty might depend on the familiarity of the research team employing the proposed method with the target technology. If researchers are unfamiliar with the focal technology, they should first conduct a broader exploratory analysis (Roper et al. 2011, pp. 76–77).

Situating the technology diffusion monitoring project involves the identification of earlier findings on technology diffusion, which have been generated by other researchers or from commercial sources. Such reports and results on commercial diffusion of technology might be available, for example, from trade associations, market research firms, consulting firms, national statistical offices, or press. The purpose of situating technology diffusion monitoring is twofold. First, it uncovers the level of technology diffusion reported by others. These insights enable the comparison of the results from other sources against the outputs from our method (in step 4). Second, situating the project contextualizes the understanding of the diffusion process for the target technology. This understanding includes previously used definitions and scope of technology, samples of companies, methods (particularly their shortcomings), and timeframes.

Step 2: Sampling and content retrieval

Once the project scope has been defined and situated within the context of the target technology, the next step is to narrow it down and focus. This involves the selection of target companies, as well as a suitable timeframe and text content. This step concludes with the search and retrieval of unstructured text content for the analysis in the next step.

The selection of companies included in the monitored sample is vital because company size and industry are strongly associated with the rate and level of technology diffusion (Fichman 2000; Greenhalgh et al. 2008, p. 139; Oliveira and Martins 2011). Furthermore, the type of companies to be monitored will also determine the range of unstructured text sources potentially available for the analysis. Some types of content, such as websites or press articles, might be available across a wide range of companies, while larger companies might also generate content in the form of, for example, press releases, annual reports, or transcripts of executive presentations. Another aspect of content selection is its alignment between the scope of technology monitoring and the role of the technology for target companies. For example, strategically important technologies for companies in the logistics industry are likely to be discussed by these companies in press releases or annual reports, but less so in the same types of content coming from the healthcare industry, where the same technologies might still be applicable, but are not as important.

Determination of the relevant timeframe is also an integral part of this step. At the initiation of the monitoring project, it is necessary to decide how far back in time to go. Identifying landmark events indicating technology commercialization serves that purpose well. For example, an event indicating the commercialization of wind turbines is the first installation of a utility-scale wind turbine farm by an energy company. Alternatively, patent analysis can provide insights into when a focal technology begins to enter the commercialization stage (Porter and Cunningham 2004, pp. 284–285). If, however, the monitoring project is a rerun or update of previous research, only recent information needs to be analyzed.

This step concludes with content search and retrieval. These tasks leverage technology-related keywords identified in the previous step. The content search involves the identification of documents with unstructured text content where there are references to the target technology and companies. The execution of the content search can either rely on existing commercial and open databases or custom-built approaches for content identification and retrieval. Potentially suitable content types include annual reports of listed companies, press articles, social media postings, technical reports, "gray" literature, company websites, and transcripts of executive presentations. A more in-depth discussion of the data sources, search, and content retrieval goes beyond the scope of this paper. Other authors have covered these topics in the past (for example, see Martino 2003; Mikova and Sokolova 2019; Porter and Cunningham 2004, secs. 6–8; Roper et al. 2011, sec. 5.2).

Step 3: Analyzing and classifying content

This step relies on qualitative content coding (Saldaña 2015). In the qualitative content analysis tradition, a code is "a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute" to the section of text being analyzed (Saldaña 2015, p. 3). In our method, researchers generate the codes. This approach to qualitative content coding is in line with the provisional coding method, which utilizes a researcher-generated and predetermined list of codes used in the analysis (Saldaña 2015, pp. 120–123). Thus, the creation of the coding scheme (see Appendix 1) must precede the content analysis. The intention behind this scheme is to align it with the objective of the commercial technology diffusion monitoring project. Hence, the codes represent phases in the commercial adoption of technology by the target companies. Past research can provide a starting point for coding scheme development. For example, Rogers (2010) proposed a generic model describing the process of technology adoption by organizations. Appendix 2 presents a non-exhaustive list of models defining phases in innovation or technology adoption by organizations. Researchers employing the proposed method should select the initial coding scheme based on its suitability in the context of the technology monitoring project.

The analysis comprises four elements: (1) selection of the coding unit; (2) testing of the initial coding scheme on a subsample of the content; (3) potential rearrangement of the scheme and another round of testing; and (4) coding of the entire sample of available content. The coding or recording unit is the “unit of text to be classified” (Weber 1990, p. 22). Since the source documents usually link to only a single company, it is sufficient to assign a single code to the whole document. However, if a single document relates to more than one company, it is necessary to narrow down the coding unit to ensure an unambiguous link between codes and individual companies. Furthermore, smaller coding units, such as paragraphs, also facilitate post-processing and post hoc analysis. For example, technology use case analysis is conducted faster when leveraging paragraph-level rather than document-level coding. After the selection of the coding unit, it is possible to test the coding scheme. Since the codes are predetermined by the researchers before analyzing the content, “[t]esting not only reveals ambiguities in the rules but also often leads to insights suggesting revisions of the classification scheme” (Weber 1990, p. 24). We suggest coding randomly selected documents representing approximately 5%–10% of the overall sample to test the coding scheme. Testing should allow researchers to evaluate whether the coding scheme granularity level is suitable. Another recommendation is to initialize the scheme with a high number of technology adoption phases. Such granularity captures finer detail from the content, if available, and thus is more informative. Researchers following this procedure must also consider reliability. There are many approaches to ensure the reliability of qualitative content coding, some of which involve quantitative measures of reliability, while others restore to consensus between raters and group discussions (Saldaña 2015, pp. 27–28). Irrespective of the selected method for establishing reliability, the researchers involved in the project should transparently report it in their study.

Step 4: Presenting, exploring, and exploiting the results

The results from the previous step need to be further processed to derive insights from technology diffusion. After coding the entire sample, the results need to be aggregated on a company and code level because it is likely that a single company will be associated with multiple documents and codes. The procedure for aggregation includes two steps: (1) sorting the documents by the company and by date from oldest to the most recent; and (2) for each company-code combination, recording the earliest date in a table. The resulting table should include company names (in rows) and phases of technology adoption included in the coding scheme (in columns). The values in the table should show dates when individual companies reached specific phases of technology adoption. Some of the cells in the resulting table are likely to be blank due to no available information. This procedure assumes that if a single company is associated with a given phase of technology adoption on a particular date, then it cannot be “degraded” to an earlier phase, even if there is a code representing a lower phase of technology adoption assigned to it on a later date. For example, researchers may code company A as reaching full-scale technology adoption in January of a given year. At the same time, based on another document from December of the same year, they may assign it a code representing testing of that technology. In that case, we assume that January is when the company has reached the full-scale commercial adoption of that technology. Researchers might still use the information about technology testing from December, for instance, in post hoc analysis (outside of the present method's scope). However, it does not impact the date of commercial adoption of the technology for company A. Thus, each date in the results table represents the earliest identified record of a given company reaching a specific phase of technology adoption.

Case Study: AI diffusion among S&P 500 companies

To illustrate the proposed method, we take the case of AI diffusion among S&P 500 companies. AI is not a single technology but rather a technological "frontier of computational advancements that references human intelligence in addressing ever more complex decision-making problems" (Berente et al. 2021, p. 1435). Technologies falling under the current umbrella of AI, most notably machine learning (ML) (Berente et al. 2021), have been recognized by executives in charge of firms around the world as having the potential to significantly impact the strategic position of their organizations and the competitive dynamics of industries (Ransbotham et al. 2020). Despite the resulting interest in AI, there is still a long way for many companies and industries to go to successfully implement the technology and have a meaningful impact on business results (Benbya et al. 2020). Hence, monitoring the progress of AI adoption by companies is a good choice for presenting the proposed method, as this technology is currently in the process of commercial diffusion among firms, particularly large firms (Benbya et al. 2020). In the remainder of this section, we present the application of the proposed method in the case of AI diffusion among S&P 500 companies.

Step 1: Scoping and situating AI diffusion

To align our search keywords with this objective, we selected "artificial intelligence" as the first target keyword. Furthermore, we recognized that the meaning of AI has been changing over recent decades (Berente et al. 2021). Therefore, we needed to limit the project scope to the latest wave of AI diffusion, which we achieved by including another broad search term representative of the current wave of AI. That second target keyword was "machine learning." Since we were interested in diffusion across all sectors, we did not want to favor any specific AI use case or application. Thus, we decided not to include any narrower keywords. For this method demonstration, we considered "artificial intelligence" and "machine learning" sufficient keywords to capture AI diffusion among companies.

Situating AI diffusion monitoring also involved the identification of past research and other reports on the topic. We investigated three types of sources: (1) academic research, (2) national statistical offices and other governmental or not-for-profit organizations, and (3) consulting firms and other commercially oriented organizations publishing such findings. We present an overview of key findings from each of these sources in the remainder of this subsection. Academic research concerned with or related to the diffusion of AI technologies has been expanding rapidly due to many new challenges and opportunities presented by AI (Benbya et al. 2021). Despite this interest, based on our review of the literature, scholars have largely overlooked the question of the level of AI diffusion among companies; thus, monitoring the commercial diffusion of AI has not been a focus. We have identified only a few studies that at least partially attempted to do that. In a study (Lyu and Liu 2021) investigating keywords related to AI and other technologies in job postings made by energy firms between 2010 and 2019, AI was the most common technology. It appeared in the content of 4%–8% of job postings, depending on the year. Another study (Weber and Schütte 2019) investigating AI adoption by ten globally leading retail companies analyzed content from publicly available sources generated by these companies and the press. The results indicate that eight out of ten companies leveraged AI, although there were significant differences in the level of AI infusion into the daily business operations of these companies. Finally, an annual AI Index Report (Daniel Zhang et al. 2021, p. 106) provides the absolute number of "AI" and "machine learning" mentions in corporate earnings calls. These numbers (nearly 5,000 and 1,400 mentions for AI and ML, respectively) can be compared against the historical peak of slightly above 5,000 and 2,000 mentions, respectively, and the mention counts for other technologies (which had significantly lower counts). In contrast to the limited number of studies related to AI diffusion monitoring, research giving insights into the determinants and process of AI adoption by individuals and organizations, as well as the antecedents and consequences, has been flourishing (van den Broek et al. 2020, 2021; for example, see: Grønsund and Aanestad 2020; Lebovitz et al. 2021; Lou and Wu 2021; Mayer et al. 2020; Reis et al. 2020; Strich et al. 2021; Dan Zhang et al. 2021; Zhang et al. 2020). These studies provided rich contextualization for this technology monitoring project and can inform exploration and interpretation of the results.

Understanding the diffusion of AI into commercial use by companies has been high on the agenda of many national statistical offices, government-related entities, and other not-for-profit organizations. The high priority of this topic results from the potentially high impact of AI on the economy (Ransbotham et al. 2020). The first finding that is prevalent across the results from different countries and institutions conducting surveys is that the overall level of AI adoption is relatively low, ranging between 1% and 20% (Eurostat 2020; Montagnier and Ek 2021; Zolas et al. 2020, p. 12). Next, large organizations generally have higher adoption rates of AI than small and medium enterprises (Eurostat 2020; Montagnier and Ek 2021; Zolas et al. 2020, p. 12). There are, however, significant differences between countries. For example, the share of large enterprises with over 250 employees that analyze big data internally using machine learning is 41% for Ireland and less than 5% for countries such as Cyprus, Lithuania, Bosnia and Herzegovina (Eurostat 2020). These findings come from surveys, which suffer from limitations beyond those we discussed previously. For example, different national statistical offices rely on their own definitions of AI, thus limiting the comparability of the findings (Montagnier and Ek 2021). Some studies include multiple technologies in a basket, thus limiting the visibility of AI-only diffusion. Notably, some studies in this category employ methodologies other than surveys, namely, content analysis of company websites (Mattila et al. 2017) and patent analysis (Toole et al. 2020).

Finally, management consulting firms and other commercially oriented organizations have been the most active publishers of reports on the state of AI diffusion among companies. These reports represent the majority of the volume and variety of insights on AI diffusion out of the three types of sources we have identified. Given the sheer number of publications in this category, we concentrated on a selected few, which we considered the most representative, informative, and credible. This selectiveness means that we left out many of the reports falling into this category. We justify this decision with the significant limitations faced by publications of this type (Montagnier and Ek 2021). Frequently, the methods used were not transparent or, at least, not replicable. Since some studies sourced survey responses from proprietary contact lists (neither random nor theoretical sampling), which were undisclosed for commercial and confidentiality reasons, they were not accessible to impartial third parties. Thus, such studies were not replicable, even if they provided generic sample descriptions. These practices might lead to (un)intentional selection bias by targeting, for example, (prospective) customers with survey questionnaires. We also excluded from our analysis some reports that intentionally introduced selection bias by targeting only respondents from firms already engaged in AI activities. These reports ignored companies to which AI has not yet diffused. Finally, the commercial interests of the report writers may conflict with their readers' interests. On the positive side, these reports typically went beyond covering the state of AI diffusion and investigated topics such as related challenges faced by organizations, level of in-house expertise, numbers, type, budget, and importance of projects related to AI, roles, and count of employees involved in AI. Additionally, these reports tended to be more up-to-date than the results from academic publications or national statistical offices, given their publication volume and frequency. Overall, these reports provided us with rich insight but required careful consideration of their methods and validity. We found that the level of AI adoption grew steadily from 2017 to 2020, with commercial AI adoption reaching 50%–60% of survey respondents or companies surveyed (Balakrishnan et al. 2020; Bughin et al. 2017; Cam et al. 2019; Chui and Malhotra 2018; Lorica and Loukides 2018; Lorica and Nathan 2019; Magoulas and Swoyer 2020; Ransbotham et al. 2017, 2018, 2019, 2020). Thus, the process of commercial diffusion of AI is still underway as we write this paper.

Step 2: Focusing on S&P 500 companies and the current wave of AI commercialization

To further narrow down the scope of the technology monitoring project, we decided to concentrate on the largest companies in a single country. We selected the largest U.S.-based companies as our target population. Based on the findings from the previous step, they were among the most advanced users of AI. Furthermore, the choice of a single country increased the homogeneity of sample companies and the content to be analyzed. These companies share an external environment and present similar internal institutional characteristics. This setting makes them sufficiently comparable to jointly analyze their commercial diffusion pattern for AI. Next, all these companies produce content in English, which allowed us to carry out the analysis in a single language only. We assumed that companies included in the S&P 500 index were representative of the target population.

Selecting the specific timeframe to be used in the analysis was the next task. Since AI has been changing the meaning over time, we wanted to exclude earlier waves of AI from the timeframe. The technology category representing the earlier wave was "expert systems" (Berente et al. 2021). We used that keyword and searched in the Scopus database for academic papers mentioning it to identify that wave. The number of articles including "expert systems" in the title, abstract, or keywords stabilized after approximately year 2000. Next, we identified the timing of several landmark events, which coincided with the start of the current commercialization wave of AI. Such events include, among others, the use of GPUs (graphics processing units) to train artificial neural networks for the first time by Andrew Ng in 2009; IBM Watson winning in Jeopardy in 2011; deep neural network-based algorithm winning the ImageNet image classification contest in 2012; and Google's AlphaGo winning against Lee Sedol in the game of Go in 2016 (Chui et al. 2018). We decided to fix the start of the technology monitoring timeframe to January 2004, which gave five years before the first identified landmark event from the current wave of AI and four years after the number of papers related to "expert systems" stabilized. The end of the monitoring timeframe coincided with the date we retrieved the data, which was the end of May 2019.

Next, we selected the content for analysis in the technology monitoring project. Based on the findings from earlier research and reports covering AI diffusion and use by companies, we knew that the technologies in our scope were of strategic importance. Thus, we decided to use transcripts of quarterly earnings calls and other investor presentations as content for the analysis. All sample companies were publicly listed, which meant that they all produced this type of content. Since investor events typically take the form of online conferences, detailed transcripts were available. Such events are the hallmark of voluntary disclosure (Rogers 2000) and serve two primary purposes for firms: informational and relational (Crawford Camiciottoli 2010). Tasker (1998) found that companies that provide less informative financial statements tend to make up for it with increased information content in conference calls. Additionally, the information content of the conference calls typically goes beyond the financial figures and includes forecasting and discussions on future trends, other relevant topics, and an unscripted Q&A session (Crawford Camiciottoli 2010). Thus, some investor calls include a discussion on technology development and adoption by companies. This type of content is not without limitations, such as evidence that executives engage in promotional rhetoric aimed at instilling investor confidence (Crawford Camiciottoli 2010) and may make deceptive statements (Larcker and Zakolyukina 2012). Executives might also not disclose the use of strategically important technologies. This secrecy may originate from the fiduciary responsibilities they hold toward the corporations employing them (Tiwari and Ahamed 2018) and, in some cases, personal liability. Despite these limitations, some scholars have utilized such transcripts as input data for their analysis. For example, Wang and colleagues (2020) used transcripts of earnings calls in connection with an ML-based personality trait detector to analyze executive personality impact on mergers and acquisition intensity. Teece (1980) used transcripts of speeches by corporate officials, in combination with other content, to study the diffusion of administrative innovation among large U.S. firms. Based on these findings, we concluded that transcripts of earnings calls and other investor presentations had the potential to be a suitable source of unstructured text for this method demonstration.

We retrieved 2,047 investor event transcripts of S&P 500 company executive presentations from the Thomson Reuters Eikon database. The search query was case insensitive and was "*artificial intelligence*" or "*machine learning*". We included only events that took place between January 2004 and May 2019. Furthermore, these events were limited to quarterly earnings calls, conferences, financial analyst days, and other investor events targeting the business and investor community. The transcripts were in raw text (unstructured) format and included three metadata fields: event date, RIC (company identifier used in the database), and company name. Additionally, we collected from the same source the following data on each sample company and based on the latest available full financial year: annual revenue, primary and secondary NAICS sector codes and the respective sector names, yearly revenue per sector code (where available), and company sector based on the assignment to S&P sector indices. We used these additional data (in step 4) for the exploration and validation of the results.

Step 3: Analysis of transcripts and classification of companies into three phases of AI implementation

Before performing the analysis of content, we initialized the coding scheme based on past research. Subsequently, we tested it on a subsample of the content and revised iteratively until concluding the process with three codes: (1) mentioning AI; (2) piloting AI; and (3) commercial use of AI. Table 1 provides definitions of the codes and examples of quotes illustrating the type of statements made by company executives, which led us to assign these codes. Next, we describe in greater detail the procedure of the coding scheme development.

We initialized the coding scheme development by considering a well-known model of the information technology implementation process (Cooper and Zmud

Table 1 The final coding scheme used in the analysis of executive presentation transcripts

Code	Definition of the code	Examples from coded texts
1: Mentioning AI	– Reference to specific plans regarding AI or ML technology implementation – Expression of interest in or intention to implement the technology in the future – Other general reference to AI or ML	– "And to the extent that we can get machine learning on the volume of data that we collect, I think that's a great opportunity for us." – "As you would expect, head count additions primarily align with our priority areas, such as cloud and apps and machine learning."
2: Piloting AI	– Reference to ongoing tests, trials or experiments that involve AI or ML technologies – Any implementation of the technology that is not yet used in regular business (not part of product/offering nor regular business process) and no information regarding timing of commercial use – AI or ML related acquisition or partnership with no details on degree of commercial use of AI or ML	– "We're doing a lot of work in our labs looking ahead again to the next few years in things like […] artificial intelligence which is moving very fast […]." – "The acquisition that we have now […], a small company but really brings some great machine learning and vision tools […]." – "[…] machine learning, we are […] really prototyping that technology internally […]"
3: Commercial use of AI	– Reference to a current commercial use of AI or ML technologies (as part of customer offering or internal processes, which are "business as usual") – Commercial launch or implementation utilizing AI or ML technologies in the near future (specific details provided)	– "We're using software and algorithms to make decisions rather than people […], especially as we insert machine learning into those decisions." – "We have had great success using […] machine-learning technologies drive those rigs to even higher levels of efficiency."

1990) and a classification scheme used in a practitioner-focused study investigating business adoption of AI (Ransbotham et al. 2017). The former model includes six phases: (1) initiation, (2) adoption, (3) adaptation, (4) acceptance, (5) routinization, and (6) infusion, while the latter includes five classes: (1) has not adopted AI and has no plans to do so, (2) has not adopted AI but plans to do so in the future, (3) has one or more AI pilot projects, (4) AI is incorporated in some processes and offerings, and (5) AI is extensively incorporated in processes and offerings. The two schemes present a high degree of alignment with each other. Given the focus of the latter on AI, we decided to use that scheme as our initial codes, with the addition of one code—other nonbusiness-related references to AI—to account for executives referring to AI or ML in a general sense or without giving sufficient detail on the level of commercial adoption. This initial coding scheme had a high degree of granularity and, thus, could capture a great degree of nuance in the data, if available.

Since the exclusive focus of our analysis was the identification of AI adoption phases by companies, we selected the unit of content analysis to be an individual transcript. The code assigned to each analyzed transcript corresponded to the highest degree of technology adoption identified within that transcript. This coding unit was deemed sufficient to meet the objectives of this method demonstration. Furthermore, this approach allowed us to focus the qualitative analysis only on those parts of the transcripts related to AI or ML mentions. This approach meant that for each transcript, we first identified all occurrences of relevant keywords and iteratively read paragraphs surrounding these keywords to determine sections of text that were relevant for the analysis and provided sufficient context to classify that individual document. If more than one section of text included references to AI, we coded the transcript with the highest identified level of AI adoption.

Two researchers (the first author and a research assistant knowledgeable about business use of information technology) tested the initial coding scheme. We independently coded 100 randomly selected transcripts, which represented approximately 5% of our document sample. After cross-checking the results and discussing whether the codes captured the relevant information in the transcripts, we concluded that there was a need to reduce the granularity of the coding scheme; thus, we lowered their number to four. After another round of coding, which included another set of 100 randomly selected transcripts, we cross-checked and revised the coding scheme again. The final coding scheme emerged, consisting of three codes (see Table 3). While the revision of the coding scheme aligned it better with the underlying data, this came at the expense of lower granularity, especially in the latter phases of AI adoption. This reduction in granularity points to potential limitations regarding executive transcripts as the sole content source for comprehensive technology diffusion monitoring. Nevertheless, it did not prevent us from demonstrating the proposed method and generating new insights in this case study of AI diffusion.

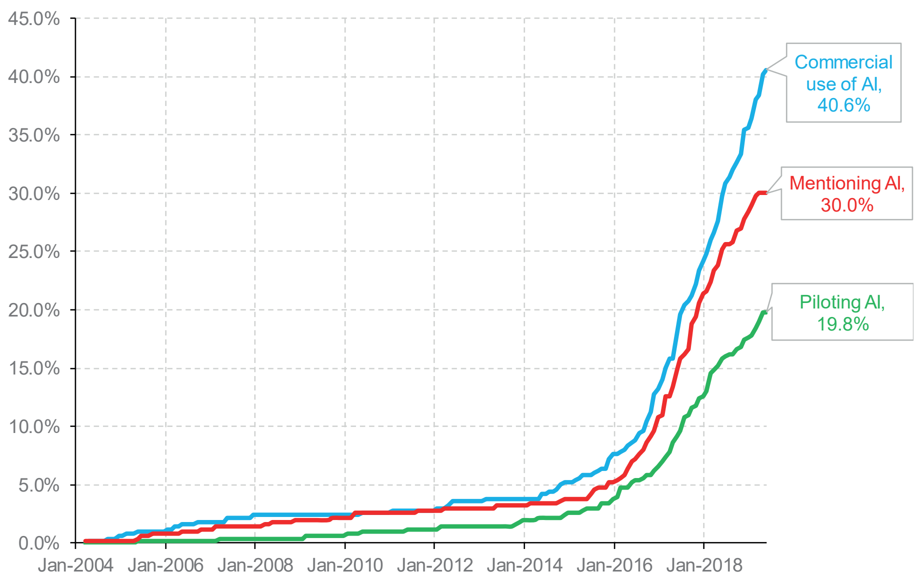
After we coded all transcripts independently, we cross-checked the results, and any differences in codes were revised and finalized through a consensus decision. According to Saldaña (2015, p. 28), this is one of the approaches used in qualitative analysis to improve consistency and address the discrepancies between coders. If our document sample was much larger, thus making parallel coding and cross-checking of the results unfeasible, or if we relied on more coders, we would restore to quantitative reliability measures.

Step 4: AI diffusion among S&P 500 companies

By aggregating the results from the previous step, which were on a document level, we arrived at the final results representing the phases of AI commercial adoption by individual S&P 500 companies throughout the monitoring timeframe. A total of 62.2% of the sample companies were assigned at least one code by the end of the study's timeframe (May 2019). As presented in Figure 2, the cumulative percentages of sample companies that reached commercial use of AI, piloted AI, and mentioned AI during investor events were 40.6%, 19.8%, and 30%, respectively.

These results are not in line with the expected sequence of technology awareness, which is followed by piloting and, later, commercial use. In other words, we expected the blue curve representing commercial use of AI to be below the two curves and

Figure 2 Cumulative percentage of companies by AI adoption phase



not above them. Based on these results, more companies reported commercial use of AI than those that either piloted AI or mentioned it in general terms during investor events. We interpret these findings as evidence of corporate executives being reluctant to build expectations by disclosing piloting of AI or referencing AI developments when their company has limited visibility on commercial implementation of AI. We conclude that the results understate the actual percentage of companies aware of AI or piloting AI. Therefore, in our subsequent discussion we primarily rely on the estimates relating to the commercial use of AI.

Validation of the results

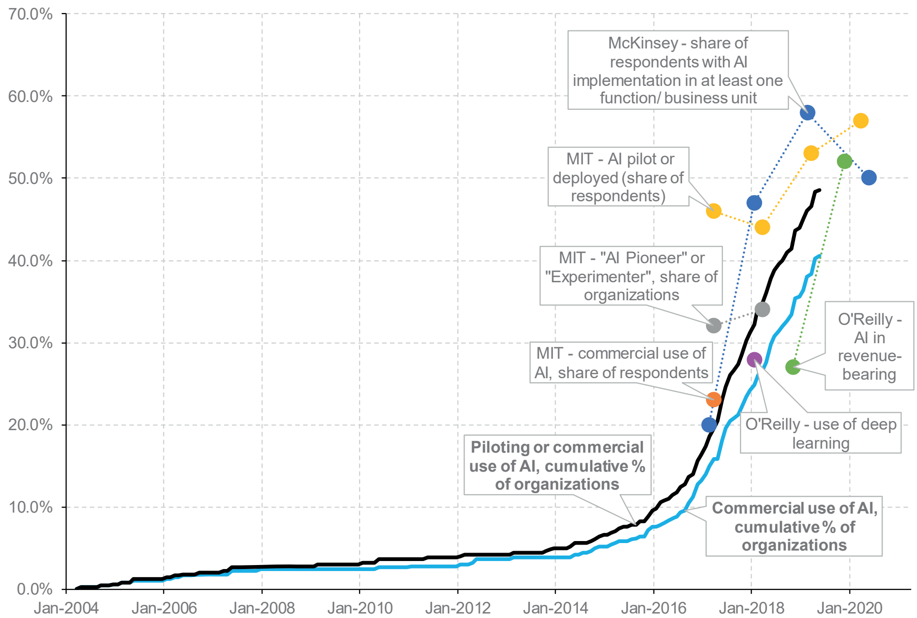
The researchers and practitioners employing the proposed method could pursue different ways of further exploring and exploiting the results presented in the previous section. We use the case study of the commercial diffusion of AI among S&P 500 companies to validate the method. We do that by comparing our results against survey-based empirical findings and two theory-based hypotheses. In this case study, our results are consistent with both empirical findings and theory.

Comparison with survey-based AI diffusion estimates

In this subsection, we compare our results on AI diffusion among S&P 500 companies with the results from several longitudinal surveys of AI use by companies, which we have identified in step 1 of the procedure. We recognize that the empirical results from these surveys are not necessarily directly comparable with our results. Thus, no formal tests can be applied here. We rely on visual inspection of Figure 3 in the results as a means of validation.

Despite limited comparability, our results on AI use by companies present an overall agreement with the trends indicated in the surveys. The alignment applies to both levels and timing. Based on this consistency, we conclude that our method provides a similar level of insight into the state of technology adoption as do commercially generated surveys. Our method, however, presents several advantages over these surveys. First, the proposed approach is transparent because it relates to a clearly defined sample of companies. The method results are also replicable due to an explicitly defined coding scheme and rules for content analysis. Another advantage of the proposed method is that the results capture a longitudinal progression of the diffusion trajectory with high granularity. What follows from transparency, replicability, and granularity is the ease of updating the results in synch with the availability of new content. Thus, the proposed method does not suffer from long time lags, which is the case with surveys. Based on the case study of AI diffusion, there appears to be

Figure 3 Comparison of the cumulative percentage of companies reaching commercial use of AI (solid blue line) and commercial use or piloting of AI (solid black line) generated using the proposed method and the results from multiple longitudinal surveys on AI use by companies



Sources of survey results: McKinsey (Balakrishnan et al. 2020; Bughin et al. 2017; Cam et al. 2019; Chui and Malhotra 2018); MIT (Ransbotham et al. 2017, 2018, 2019, 2020); O'Reilly (Lorica and Loukides 2018; Lorica and Nathan 2019; Magoulas and Swoyer 2020).

no qualitatively significant difference between the reported levels of AI use in surveys and those generated based on executive presentations geared toward investors. Thus, our method shows that it is possible to gain insight into the commercial diffusion of technology without privileged access to information using, for example, publicly available investor presentations. This result is relevant to practitioners who do not have information access similar to that of management consulting companies or other commercial organizations carrying out market analysis.

Investigation of differences between sectors

Next, we validate the results by comparing the outputs of the proposed method with the predictions generated from theory. Based on innovation diffusion theory and research results on information technology diffusion among organizations, we expect that there should be meaningful differences in the rate of AI diffusion between companies from different sectors (Fichman 2000; Greenhalgh et al. 2008, p. 139; Oliveira and Martins 2011). Therefore, we can state the first null hypothesis as follows:

H1: There is no difference in the commercial diffusion rate of AI between companies from different sectors.

We can statistically test the difference between sector-level diffusion rates by investigating stochastic dominance between the diffusion curves for each sector. We examine stochastic dominance using the Kruskal–Wallis (KW) test, which is non-parametric and suitable for testing multiple groups at once (Mangiafico 2016, pp. 248–261). The test rejects the null hypothesis that there is no stochastic dominance between any pair of sectors (see Table 2). Next, we perform a post hoc analysis using the Dunn test for pairwise comparison to determine stochastic dominance individually between each pair of sectors (Mangiafico 2016, pp. 255–256). Based, on this we conclude that the IT, financial, communication services, and healthcare sectors implemented AI into commercial use significantly earlier than companies in the real estate, materials, and utility sectors. These results are consistent with expectations and past empirical findings (Fichman 2000). Despite all S&P 500 companies being large corporations based in the U.S., there are meaningful differences between their commercial adoption rates of AI. In sectors where competitive pressures are highest and innovation is a driver of success, commercial AI adoption is significantly higher than in traditional sectors where fixed assets are the determinant of business success. Consequently, these results provide a validation of the proposed method against the theory.

Table 2 Results from the Kruskal–Wallis test for stochastic equality between the timing of commercial adoption of AI by different sectors

chi-squared	df	p-value
87.85	10	< 0.001

Investigation of differences between digital intensity levels

We perform another validation of the results from our methods by comparing the diffusion rates for commercial use of AI between companies exhibiting different levels of related knowledge. Related knowledge is one of the determinants that drive the adoption of information technologies by organizations (Fichman 2000; Greenhalgh et al. 2008, p. 12; Pennings and Harianto 1992). We operationalize related knowledge through the measure of the digital intensity of a firm. Digital intensity is a multifaceted indicator of how much firms "went digital" (Calvino et al. 2018). It measures the adoption of advanced digital technologies, employing human capital skilled with these technologies, and the extent of leveraging digital tools in relationships with customers and suppliers (Calvino et al. 2018). Based on recent empirical evidence, digital intensity is associated with AI adoption (Kinkel et al. 2021; Radhakrishnan and Chattopadhyay 2020). We use a method for approximating the digital intensity of a company based on aggregated measures of industry-level digital intensity and the level of firm engagement in different industries (Mucha and Seppälä 2021). We measure this engagement using revenue derived from activities recorded under individual business units of a company (Mucha and Seppälä 2021). Since these business units are associated with industry codes, we can map their industry-level digital intensities to the firm level. We can state the second null hypothesis as follows:

H2: There is no difference in the commercial diffusion rate of AI between companies with different levels of digital intensity.

Based on the Kruskal–Wallis (KW) test results (Table 3), we reject the null hypothesis that there is no stochastic dominance between companies from different levels of digital intensity. Post hoc analysis based on the Dunn test reveals stochastic dominance between each pair of digital intensity levels. Based on these results, we conclude that the commercial adoption of AI is strongly associated with the firm's digital intensity level. These results are consistent with past empirical findings (Kinkel et al. 2021; Radhakrishnan and Chattopadhyay 2020), thus providing validation for the proposed method.

Table 3 Results from the Kruskal–Wallis test for stochastic equality between the timing of commercial adoption of AI by companies with different digital intensity levels

chi-squared	df	p-value
54.31	2	< 0.001

Discussion and concluding remarks

Technology monitoring is the process of observing and keeping up with developments in a specific technology (Roper et al. 2011, p. 72). It is critical to R&D management, technology management, and overall business strategy (Burgelman et al. 2004, pp. 8–9). Thus, scholars and practitioners frequently rely on technology monitoring to generate new insights and knowledge. However, the predominant focus for the development of new methods for technology monitoring has been on patent analysis or otherwise precommercial stages of the technology lifecycle. These developments resulted in national statistical offices and commercial organizations relying on crude methods for monitoring the commercial diffusion of technologies, such as survey-based research developed in the 1940s and 1950s. In this paper, we propose an alternative approach to monitoring the commercial diffusion of technology.

The proposed method builds on past research within the technology monitoring and innovation diffusion literature. By utilizing qualitative content analysis, while following the procedure we propose, it is possible to generate high granularity time series representing the diffusion of technologies from early phases of commercial adoption, such as awareness of technology, to commercial use. This analysis leverages unstructured text, which can take different forms, such as the text of websites, press articles, press releases, annual reports, or transcripts of executive presentations.

We illustrate the proposed method by analyzing the commercial diffusion of AI technologies among S&P 500 companies during the January 2004–May 2019 period using 2,047 transcripts of quarterly earnings calls and other investor events. Based on qualitative content analysis of these transcripts, we assign them to one of three groups: (1) mentioning AI during investor events, (2) piloting AI, or (3) using AI in a commercial context. We find that by the end of May 2019, 40.6% of companies had reached the commercial use phase of AI, 8% reported piloting AI, and 13.6% mentioned AI in general terms only. We conclude the analysis by carrying out a validation against existing empirical findings on AI use by companies and theoretical predictions derived from the research on the diffusion of information technology among organizations. The results align well with survey results on AI diffusion published by management consulting firms and other commercially oriented organizations. Unlike these surveys, however, our method is transparent, replicable, and does not require privileged access to information, as transcripts of investor events are readily available from various databases. Another advantage of our method is that its results are available without time lags commonly associated with periodic surveys. A comparison of our results with the theoretical predictions shows consistency between the two. Our results on the differences in commercial diffusion rates for AI between companies from different sectors are consistent with expectations based on information technology diffusion research (Fichman 2000; Oliveira et al. 2019). Sectors where competitive pressure and innovativeness are high, such as IT, communication services, finance, and healthcare, adapted AI more rapidly than traditional sectors, such as utilities, real estate, and basic materials, whose fixed assets are the main determinants of competitiveness. Additionally, the results generated by our method showed that firms exhibiting a high level of digital intensity were faster commercial adopters of AI than medium or low digital intensity firms. This impact of related knowledge on the pace of AI adoption is consistent with past results from research on both information technology and AI by companies (Kinkel et al. 2021; Radhakrishnan and Chattopadhyay 2020). Overall, this illustration of the proposed method using the case of AI diffusion gives practically relevant insights and shows that the results are consistent with both past empirical findings and theoretical predictions.

This paper contributes to IS research concerned with technology monitoring and innovation diffusion as well as to practice. First, despite commercial adoption of technologies being essential to their generation of impact on economy and society (Hall 2004), this latter part of the technology life cycle has been grossly overlooked by researchers developing methods for technology monitoring. Our approach to monitoring addresses this gap by targeting technologies that enter the commercialization stage of their lifecycle or are in widespread use. Second, the proposed method is generally applicable to a wide range of technologies and contexts. This versatility results from reliance on unstructured text content as data input and broadly defined phases of technology adoption at the initiation of the analysis. Researchers employing the proposed method can fine-tune the specific content type and granularity of technology adoption phases to fit their research context. This broad applicability of the method means that it can be incorporated into and enrich a variety of studies investigating topics related to technology diffusion and adoption by organizations. These studies typically rely on surveys alone for data collection. Thus, they could increase robustness by triangulating some of the results with the method we propose. For practitioners, our method presents a transparent, replicable, and updatable alternative to commercially run surveys. Given that the proposed approach is longitudinal, ongoing technology monitoring activities carried out by strategy teams inside organizations can benefit from only incremental efforts needed to update the results with the latest analysis.

This research exhibits certain limitations and presents opportunities for further development. Since the proposed method relies on unstructured content analysis, it is of limited utility for analyzing technology diffusion among companies that generate little or no such content, such as some subpopulations of early-stage start-ups and small- and medium-sized enterprises. Next, this method might underperform surveys for studies of technologies that are of low importance to target companies. Even if the target companies generate unstructured text, mentions of such technologies might be absent there. Finally, this research relied exclusively on the case of AI diffusion among S&P 500 companies to validate the method's performance. Future validation should include a broader range of technologies, as well as types of unstructured content. Furthermore, given the continued advancements in natural language processing methods and ML, in general, the proposed method could serve as the foundation for an automated technology monitoring algorithm or tool for monitoring the commercial diffusion of technology. Such future advancement would resemble the development path of methods used in technology monitoring based on patent analysis.

References

Armour, H. O., and Teece, D. J. 1978. "Organizational Structure and Economic Performance: A Test of the Multidivisional Hypothesis," The Bell Journal of Economics (9:1), p. 106. (https://doi.org/10.2307/3003615).

Balakrishnan, T., Chui, M., Hall, B., and Henke, N. 2020. Global Survey: The State of AI in 2020, McKinsey.

Benbya, H., Davenport, T. H., and Pachidi, S. 2020. "Special Issue Editorial: Artificial Intelligence in Organizations: Current State and Future Opportunities.," MIS Quarterly Executive (19:4), pp. ix–xxi.

Benbya, H., Pachidi, S., and Jarvenpaa, S. 2021. "Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research," Journal of the Association for Information Systems (22:2), Atlanta, United States: Association for Information Systems, p. 10. (http://dx.doi.org/10.17705/1jais.00662).

Berente, N., Bin Gu, Recker, J., and Santhanam, R. 2021. "Managing Artificial Intelligence," MIS Quarterly (45:3), MIS Quarterly, pp. 1433–1450. (https://doi.org/10.25300/MISQ/2021/16274).

van den Broek, E., Sergeeva, A., and Huysman, M. 2020. "Hiring Algorithms: An Ethnography of Fairness in Practice," in 40th International Conference on Information Systems, ICIS 2019, Association for Information Systems, pp. 1–9.

van den Broek, E., Sergeeva, A., and Huysman, M. 2021. "When the Machine Meets the Expert: An Ethnography of Developing Ai for Hiring," MIS Quarterly (45:3), MIS Quarterly, pp. 1557–1580. (https://doi.org/10.25300/MISQ/2021/16559).

Brynjolfsson, E., Rock, D., and Syverson, C. 2018. "Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics," in The Economics of Artificial Intelligence: An Agenda, University of Chicago Press, pp. 23–57. (http://www.nber.org/chapters/c14007).

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., and Trench, M. 2017. "Artificial Intelligence: The next Digital Frontier," McKinsey Global Institute, pp. 1–80.

Burgelman, R. A., Christensen, C. M., and Wheelwright, S. C. 2004. Strategic Management of Technology and Innovation, (4th ed.), McGraw-Hill/Irwin.

Calvino, F., Criscuolo, C., Marcolin, L., and Squicciarini, M. 2018. A Taxonomy of Digital Intensive Sectors. (https://www.oecd-ilibrary.org/content/paper/f404736a-en).

Cam, A., Chui, M., and Hall, B. 2019. Global AI Survey: AI Proves Its Worth, but Few Scale Impact, McKinsey.

Chen, I., and Small, M. 1994. "Implementing Advanced Manufacturing Technology: An Integrated Planning Model," Omega (22:1), pp. 91–103. (https://doi.org/10.1016/0305-0483(94)90010-8).

Choi, H., Oh, S., Choi, S., and Yoon, J. 2018. "Innovation Topic Analysis of Technology: The Case of Augmented Reality Patents," IEEE Access (6), pp. 16119–16137.

(https://doi.org/10.1109/ACCESS.2018.2807622).
Chui, M., Kamalnath, V., and McCarthy, B. 2018. "An Executive's Guide to AI," McKinsey&Company. (https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Analytics/Our%20Insights/An%20executives%20guide%20to%20AI/An-executives-guide-to-AI.ashx).
Chui, M., and Malhotra, S. 2018. "AI Adoption Advances, but Foundational Barriers Remain," Mckinsey and Company. (https://www.mckinsey.com/featured-insights/artificial-intelligence/ai-adoption-advances-but-foundational-barriers-remain).
Cooper, R. B., and Zmud, R. W. 1990. "Information Technology Implementation Research: A Technological Diffusion Approach," Management Science (36:2), INFORMS: Institute for Operations Research, pp. 123–139. (https://doi.org/10.1287/mnsc.36.2.123).
Crawford Camiciottoli, B. 2010. "Earnings Calls: Exploring an Emerging Financial Reporting Genre," Discourse & Communication (4:4), pp. 343–359. (https://doi.org/10.1177/1750481310381681).
Cunningham, S. W., and Kwakkel, J. H. 2016. Analytics and Tech Mining for Engineering Managers, Momentum Press. (https://books.google.fi/books?id=5yKxDAEACAAJ).
Eurostat. 2020. "Eurostat - Use of Machine Learning for Big Data Analysis," Eurostat. (https://ec.europa.eu/eurostat/databrowser/view/ISOC_EB_AI__custom_1269101/default/bar?lang=en, accessed September 29, 2021).
Fichman, R. G. 2000. "The Diffusion and Assimilation of Information Technology Innovations," Framing the Domains of IT Management: Projecting the Future through the Past (105127), pp. 105–128.
Fredström, A., Wincent, J., Sjödin, D., Oghazi, P., and Parida, V. 2021. "Tracking Innovation Diffusion: AI Analysis of Large-Scale Patent Data towards an Agenda for Further Research," Technological Forecasting and Social Change (165), p. 120524. (https://doi.org/10.1016/j.techfore.2020.120524).
Grant, R. M. 2016. Contemporary Strategy Analysis: Text and Cases Edition, John Wiley & Sons.
Greenhalgh, T., Robert, G., Bate, P., Macfarlane, F., and Kyriakidou, O. 2008. Diffusion of Innovations in Health Service Organisations: A Systematic Literature Review, John Wiley & Sons.
Grønsund, T., and Aanestad, M. 2020. "Augmenting the Algorithm: Emerging Human-in-the-Loop Work Configurations," The Journal of Strategic Information Systems (29:2), Strategic Perspectives on Digital Work and Organizational Transformation, p. 101614. (https://doi.org/10.1016/j.jsis.2020.101614).
Hall, B. H. 2004. "Innovation and Diffusion," National Bureau of Economic Research.
Hong, S., Kim, J., Woo, H.-G., Kim, Y.-C., and Lee, C. 2021. "Screening Ideas in the Early Stages of Technology Development: A Word2vec and Convolutional Neural Network Approach," Technovation, p. 102407. (https://doi.org/10.1016/j.tech-

novation.2021.102407).

Kinkel, S., Baumgartner, M., and Cherubini, E. 2021. "Prerequisites for the Adoption of AI Technologies in Manufacturing – Evidence from a Worldwide Sample of Manufacturing Companies," Technovation, p. 102375. (https://doi.org/10.1016/j.technovation.2021.102375).

Larcker, D. F., and Zakolyukina, A. A. 2012. "Detecting Deceptive Discussions in Conference Calls," Journal of Accounting Research (50:2), pp. 495–540. (https://doi.org/10.1111/j.1475-679X.2012.00450.x).

Lebovitz, S., Levina, N., and Lifshitz-Assaf, H. 2021. "Is Ai Ground Truth Really True? The Dangers of Training and Evaluating Ai Tools Based on Experts' Know-What," MIS Quarterly (45:3), MIS Quarterly, pp. 1501–1525. (https://doi.org/10.25300/MISQ/2021/16564).

Lorica, B., and Loukides, M. 2018. How Companies Are Putting AI to Work Through Deep Learning, O'Reilly Media. Sebastopol, CA: O'Reilly Media.

Lorica, B., and Nathan, P. 2019. AI Adoption in the Enterprise 2019, p. 32.

Lou, B., and Wu, L. 2021. "Ai on Drugs: Can Artificial Intelligence Accelerate Drug Development? Evidence from a Large-Scale Examination of Bio-Pharma Firms," MIS Quarterly (45:3), MIS Quarterly, pp. 1451–1482. (https://doi.org/10.25300/MISQ/2021/16565).

Lyu, W., and Liu, J. 2021. "Artificial Intelligence and Emerging Digital Technologies in the Energy Sector," Applied Energy (303), p. 117615. (https://doi.org/10.1016/j.apenergy.2021.117615).

Magoulas, R., and Swoyer, S. 2020. AI Adoption in the Enterprise 2020, O'Reilly Media. Sebastopol, CA: O'Reilly Media.

Mangiafico, S. S. 2016. Summary and Analysis of Extension Program Evaluation in R, Rutgers Cooperative Extension, New Brunswick, NJ. (https://rcompanion.org/documents/RHandbookProgramEvaluation.pdf).

Martino, J. P. 2003. "A Review of Selected Recent Advances in Technological Forecasting," Technological Forecasting and Social Change (70:8), TF Highlights from ISF 2002, pp. 719–733. (https://doi.org/10.1016/S0040-1625(02)00375-X).

Mattila, J., Pajarinen, M., Rouvinen, P., and Seppälä, T. 2017. "Tekoälyn Voitto?," Digibarometri (2017), ETLA-The Research Institute of the Finnish Economy, pp. 6–30.

Mayer, A.-S., Strich, F., and Fiedler, M. 2020. "Unintended Consequences of Introducing AI Systems for Decision Making," MIS Quarterly Executive (19:4), p. 6.

Meyer, A. D., and Goes, J. B. 1988. "Organizational Assimilation of Innovations: A Multilevel Contextual Analysis," Academy of Management Journal (31:4), Academy of Management, pp. 897–923. (https://doi.org/10.2307/256344).

Mikova, N., and Sokolova, A. 2019. "Comparing Data Sources for Identifying Technology Trends," Technology Analysis & Strategic Management (31:11), Routledge, pp. 1353–1367. (https://doi.org/10.1080/09537325.2019.1614157).

Montagnier, P., and Ek, I. 2021. "AI Measurement in ICT Usage Surveys," OECD DIGITAL ECONOMY PAPERS, p. 50. (https://doi.org/10.1787/72cce754-en).

Mucha, T., and Seppälä, T. 2021. "Estimating Firm Digitalization: A Method for Disaggregating Sector-Level Digital Intensity to Firm-Level," MethodsX, p. 101233. (https://doi.org/10.1016/j.mex.2021.101233).

Oliveira, T., and Martins, M. F. 2011. "Literature Review of Information Technology Adoption Models at Firm Level," The Electronic Journal Information Systems Evaluation (14:1), p. 13.

Oliveira, T., Martins, R., Sarker, S., Thomas, M., and Popovic, A. 2019. "Understanding SaaS Adoption: The Moderating Impact of the Environment Context," International Journal of Information Management (49), pp. 1–12. (https://doi.org/10.1016/j.ijinfomgt.2019.02.009).

Oliveira, T., Thomas, M., and Espadanal, M. 2014. "Assessing the Determinants of Cloud Computing Adoption: An Analysis of the Manufacturing and Services Sectors," Information & Management (51:5), pp. 497–510. (https://doi.org/10.1016/j.im.2014.03.006).

Pennings, J. M., and Harianto, F. 1992. "The Diffusion of Technological Innovation in the Commercial Banking Industry," Strategic Management Journal (13:1), pp. 29–46. (https://doi.org/10.1002/smj.4250130104).

Porter, A. L., and Cunningham, S. W. 2004. Tech Mining: Exploiting New Technologies for Competitive Advantage, (Vol. 29), John Wiley & Sons.

Porter, A. L., and Detampel, M. J. 1995. "Technology Opportunities Analysis," Technological Forecasting and Social Change (49:3), pp. 237–255. (https://doi.org/10.1016/0040-1625(95)00022-3).

Radhakrishnan, J., and Chattopadhyay, M. 2020. "Determinants and Barriers of Artificial Intelligence Adoption – A Literature Review," in Re-Imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation, IFIP Advances in Information and Communication Technology, S. K. Sharma, Y. K. Dwivedi, B. Metri, and N. P. Rana (eds.), Cham: Springer International Publishing, pp. 89–99. (https://doi.org/10.1007/978-3-030-64849-7_9).

Ransbotham, S., Gerbert, P., Reeves, M., Kiron, D., and Spira, M. 2018. "Artificial Intelligence in Business Gets Real," MIT Sloan Management Review and Boston Consulting Group.

Ransbotham, S., Khodabandeh, S., Fehling, R., LaFountain, B., and Kiron, D. 2019. "Winning with AI," MIT Sloan Management Review (61180).

Ransbotham, S., Khodabandeh, S., Kiron, D., Candelon, F., Chui, M., and LaFountain, B. 2020. "Expanding AI's Impact With Organizational Learning," MIT Sloan Management Review and Boston Consulting Group. (https://sloanreview.mit.edu/projects/expanding-ais-impact-with-organizational-learning/).

Ransbotham, S., Kiron, D., Gerbert, P., and Reeves, M. 2017. "Reshaping Business with Artificial Intelligence: Closing the Gap between Ambition and Action," MIT

Sloan Management Review (59:1), Massachusetts Institute of Technology, Cambridge, MA.

Reis, L., Maier, C., Mattke, J., Creutzenberg, M., and Weitzel, T. 2020. "Addressing User Resistance Would Have Prevented a Healthcare AI Project Failure," MIS Quarterly Executive (19:4), INDIANA UNIV, OPER & DECISION TECHNOL DEPT, pp. 279–296.

Rogers, E. M. 1983. Diffusion of Innovations, (3rd ed.), New York : London: Free Press ; Collier Macmillan.

Rogers, E. M. 2010. Diffusion of Innovations, Simon and Schuster.

Rogers, P. S. 2000. "CEO Presentations in Conjunction with Earnings Announcements: Extending the Construct of Organizational Genre Through Competing Values Profiling and User-Needs Analysis," Management Communication Quarterly (13:3), pp. 426–485. (https://doi.org/10.1177/0893318900133003).

Roper, A. T., Cunningham, S., Porter, A. L., Mason, T. W., Rossini, F. A., and Banks, J. 2011. Forecasting and Management of Technology, (2nd ed.), Hoboken, N.J: John Wiley & Sons.

Ryan, B., and Gross, N. C. 1943. "The Diffusion of Hybrid Seed Corn in Two Iowa Communities.," Rural Sociology (8:1), Rural Sociological Society, etc., p. 15.

Saldaña, J. 2015. The Coding Manual for Qualitative Researchers, Sage.

Santos, F. M., and Eisenhardt, K. M. 2009. "Constructing Markets and Shaping Boundaries: Entrepreneurial Power in Nascent Fields," Academy of Management Journal (52:4), Academy of Management Briarcliff Manor, NY, pp. 643–671.

Segev, A., Jung, S., and Choi, S. 2015. "Analysis of Technology Trends Basedon Diverse Data Sources," IEEE Transactions on Services Computing (8:6), pp. 903–915. (https://doi.org/10.1109/TSC.2014.2338855).

Strich, F., Mayer, A.-S., and Fiedler, M. 2021. "What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity," Journal of the Association for Information Systems (22:2), Atlanta, United States: Association for Information Systems, p. 9. (http://dx.doi.org/10.17705/1jais.00663).

Tasker, S. C. 1998. "Bridging the Information Gap: Quarterly Conference Calls as a Medium for Voluntary Disclosure," Review of Accounting Studies (3:1/2), Springer Nature, pp. 137–167. (https://doi.org/10.1023/A:1009684502135).

Teece, D. J. 1980. "The Diffusion of an Administrative Innovation," Management Science (26:5), INFORMS, pp. 464–470.

Teichert, T., and Mittermayer, M.-A. 2002. "Text Mining for Technology Monitoring," in IEEE International Engineering Management Conference (Vol. 2), , August, pp. 596–601 vol.2. (https://doi.org/10.1109/IEMC.2002.1038503).

Tiwari, A. K., and Ahamed, N. 2018. "Executive Tenure and Firm Performance: An Empirical Examination of Indian Corporate Landscape *," Advances in Decision Sciences (22), Taichung, China: Asia University, Taiwan, pp. 1–29.

Toledo, C. 2005. "A Five-Stage Model of Computer Technology Infusion into Teacher Education Curriculum," Contemporary Issues in Technology and Teacher Education (5:2), Society for Information Technology & Teacher Education, pp. 177–191.

Toole, A., Pairolero, N., Giczy, A., Forman, J., Pulliam, C., Such, M., Chaki, K., Orange, D., Homescu, A. T., and Frumkin, J. 2020. Inventing AI: Tracing the Diffusion of Artificial Intelligence with US Patents, October.

Valente, T. W., and Rogers, E. M. 1995. "The Origins and Development of the Diffusion of Innovations Paradigm as an Example of Scientific Growth," Science Communication (16:3), pp. 242–273. (https://doi.org/10.1177/1075547095016003002).

Wang, Q., Lau, R. Y. K., and Yang, K. 2020. "Does the Interplay between the Personality Traits of CEOs and CFOs Influence Corporate Mergers and Acquisitions Intensity? An Econometric Analysis with Machine Learning-Based Constructs," Decision Support Systems (139), p. 113424. (https://doi.org/10.1016/j.dss.2020.113424).

Weber, F. D., and Schütte, R. 2019. "State-of-the-Art and Adoption of Artificial Intelligence in Retailing," Digital Policy, Regulation and Governance (21:3), Emerald Publishing Limited, pp. 264–279. (https://doi.org/10.1108/DPRG-09-2018-0050).

Weber, R. P. 1990. Basic Content Analysis, Sage.

Zhang, Daniel, Mishra, S., Brynjolfsson, E., Etchemendy, J., Ganguli, D., Grosz, B., Lyons, T., Manyika, J., Niebles, J. C., and Sellitto, M. 2021. "The AI Index 2021 Annual Report," ArXiv Preprint ArXiv:2103.06312.

Zhang, Dan, Pee, L. G., and Cui, L. 2021. "Artificial Intelligence in E-Commerce Fulfillment: A Case Study of Resource Orchestration at Alibaba's Smart Warehouse," International Journal of Information Management (57), p. 102304. (https://doi.org/10.1016/j.ijinfomgt.2020.102304).

Zhang, Z., Hummel, J. T., Nandhakumar, J., and Waardenburg, L. 2020. "Addressing the Key Challenges of Developing Machine Learning AI Systems for Knowledge-Intensive Work.," MIS Quarterly Executive (19:4), pp. 221–238. (https://doi.org/10.17705/2msqe.00035).

Zolas, N., Kroff, Z., Brynjolfsson, E., McElheran, K., Beede, D., Buffington, C., Goldschlag, N., Foster, L., and Dinlersoz, E. 2020. Measuring Technology Adoption in Enterprise-Level Surveys: The Annual Business Survey, January 3.

Appendix 1

Table 4 A generic structure of a content coding scheme

Code	Description of the code	Examples from texts coded
Code 1 (Earliest considered phase of technology adoption)	– Provide a description of the code. – Use of negative examples (what not to include) is also useful.	– Provide examples (quotes) that illustrate text that should be assigned that code. – Examples might not be available in the first iteration of the coding scheme development. Therefore, in the first round of coding, coders need to rely on the code definitions alone. – This column should be populated for the subsequent coding rounds.
...	...	...
Code N (Latest considered phase of technology adoption)	...	...

Appendix 2

Table 5 A non-exhaustive selection of models defining phases in technology (innovation) adoption or implementation by organizations

Source	Phases in technology (innovation) adoption or implementation by organizations
(Cooper and Zmud 1990)	1) Initiation; 2) Adoption; 3) Adaptation; 4) Acceptance; 5) Routinization; 6) Infusion
(Rogers 2010)	1) Knowledge; 2) Persuasion; 3) Decision; 4) Implementation; 5) Confirmation
(Meyer and Goes 1988)	**Knowledge-awareness stage:** 1) Apprehension: individuals learn of the innovation's existence; 2) Consideration: individuals consider the innovation's suitability for their organization; 3) Discussion: individuals engage in conversations concerning adoption. **Evaluation-choice stage:** 1) Acquisition proposal: it is formally proposed to purchase the equipment that embodies the innovation; 2) Medical–fiscal evaluation: medical and financial costs and benefits are weighed up; 3) Political–strategic evaluation: political and strategic costs and benefits are weighed up. **Adoption-implementation stage:** 1) Trial: the equipment is purchased but still under trial evaluation; 2) Acceptance: the equipment becomes well accepted and frequently used; 3) Expansion: the equipment is expanded or upgraded.
(Toledo 2005)	1) Pre-integration; 2) Transition; 3) Development; 4) Expansion; 5) Systemwide Integration

ARTICLE 3

What Have We Learned About Machine Learning?

A Meta Analysis

Tomasz Mucha[a]
Timo Seppälä[a, b]
Robin Gustafsson[a]

Affiliations:

[a] Aalto University, Finland
[b] ETLA Economic Research, Finland

Original publication:

Mucha Tomasz, Seppälä Timo and Gustafsson Robin, "Yksitoista oppia koneoppimisen meta-analyysistä" (2021), Business Finland Brief No. 4/2021.

Abstract

Recent advances in machine learning (ML) have triggered many firms to try putting the technology into commercial use. However, the creation of ML-based organizational capabilities remains a major challenge. With the aim of extending our understanding of organizational capabilities, this paper takes a socio-technical system perspective on the microfoundations of capabilities, develops an integrative conceptual framework, and discusses the resulting insights relevant to organizational ML initiatives. In contrast to past IS research, our framework is more general and versatile, since it is not restricted to dynamic capabilities only, as well as incorporates a temporal dimension facilitating the inspection of processes leading to the formation and change of organizational capabilities. This is illustrated with multiple propositions, which we develop by applying the framework to the context of organizational ML initiatives. Conceptual insights are backed with rich anecdotal evidence.

Keywords

Artificial intelligence, Machine learning meta analysis, Socio-technical systems, Organization

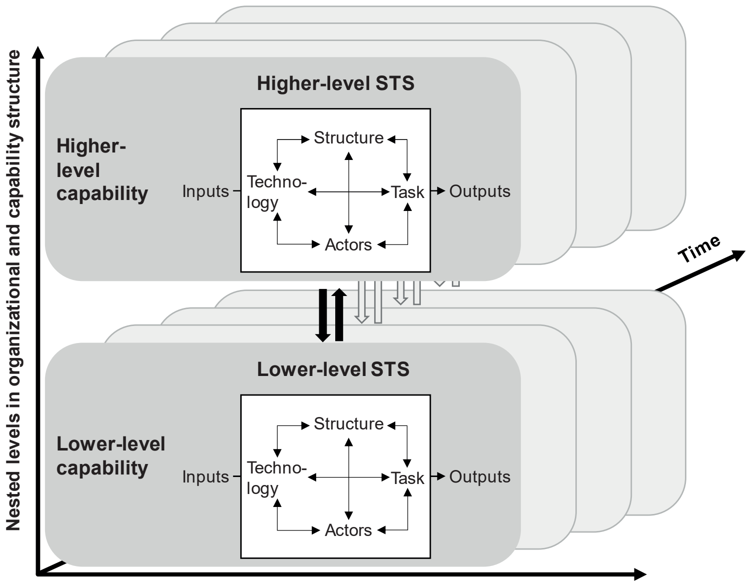
1 Introduction and conceptual framework

ML, which is at the core of the ongoing commercialization wave of artificial intelligence (AI), is currently viewed by large organizations as the most important and disruptive new technology [1]. Unlike traditional approach to programming, ML is a category of statistical and computational techniques to learning patterns and constructing inductive inferences from data or experience [1], [2]. ML coupled with modern computing resources and abundant data have enabled computers to significantly improve state-of-the-art performance on many tasks, including machine translation, speech recognition, image recognition, and generation of text, audio, and images [3]. These advances have triggered an increasing number of companies to try putting ML into commercial use [4], [5]. Despite this heightened interest, organizations face major challenges in enhancing existing or developing new capabilities with ML. While piloting ML is relatively easy, scaling and deployment have proven challenging [1], [5]. Ransbotham and colleagues [4] find that only one in ten companies gets meaningful value out of their ML initiatives. Clearly, creation of ML-based capabilities is a major challenge.

Organizational capabilities do not come into being solely based on capacities of a technology [6]–[9]. Instead, they require a process of practice and routinization, which aligns not only the tasks with technological tools, but also with skills and roles of people, as well as with the organizational structure and communication flows [6], [10]–[12]. Similarly, introducing new technology to modify or develop an existing capability might require a period of practice and routinization before the new level of capability performance is reached [12], [13]. Hence, to understand better the relationship between new technologies, such as ML, and organizational capabilities we need to study the process of capability formation and change, while explicitly recognizing the links between technology, tasks, people, and structure within which they operate. Thus, we need to shift the focus of our analysis to the level of microfoundations [14], [15]. Despite the calls for investigating microfoundations of organizational capabilities [14], [15], there is only a limited number of studies discussing this topic. Notable examples from IS domain [8], [16]–[18] recognize that socio-technical system (STS) perspective [19], [20] is a useful conceptual framing to understand the interplay between digital technologies and organizational capabilities.

We define the microfoundational elements of organizational capabilities and recognized that they correspond to the building blocks of an STS. Furthermore, our analysis of organizational capabilities and STS theory literature highlighted the prominence of the following characteristics shared by these perspectives: (1) routinization; (2) deep structure; and (3) nesting. Based on these, we propose an STS framework for organizational capabilities, as presented in Figure 1, and apply it in the context of organizational ML initiatives.

Figure 1 STS framework for organizational capabilities



Our generalized version of the framework includes two organizational capabilities, each at a different level. These capabilities are practiced and routinized transformations of inputs into outputs via the underlying socio-technical systems. The vertical axis represents levels in the capability structure, while horizontal axis spatially distributes inputs, STS, and outputs. For the sake of simplicity, we assume discrete time, thus the transformation of inputs into outputs is assumed to take place within a single time increment. The depth axis represents progression of these capabilities through time, which might involve changes in one or multiple elements, as well as bidirectional impacts between the capabilities (or their underlying STSs) on different levels. The proposed framework presents a simplified structure, which, can be extended to cover more than two levels in the capability structure, as well as more than one capability on each level. In the remainder of this section we apply this framework in the context of organizational ML initiatives and develop propositions based on that. In our discussion, we follow the path of an increasing organizational engagement with ML. We start with one-off uses of ML, then proceed to the development and use of ML capability; creation of ML-based capabilities; learning in and improvement of ML-based capabilities; and conclude with full automation of a capability.

In this paper, we assume an STS perspective on organizational capabilities and develop a conceptual framework integrating these two levels of analysis. Based on the extant literature, we identify strong links between STS-level microfoundations and organizational capabilities. The resulting framework not only captures the temporal aspect of capability change, but also interrelationships between multiple capabilities within a single organization. We then use the framework to develop insights in the context of organizational ML initiatives. In our discussion, we follow the path of an increasing organizational engagement with ML, starting from one-off uses of ML and ending with full automation of a capability. Based on these, we develop multiple propositions. The propositions are backed by and clarified with anecdotal evidence collected from published case studies focusing on organizational use of ML, as well as from an ongoing 2-year-long research involving interviews and participatory observation of a national government funded accelerator promoting and facilitating AI use in organizations (Accelerator name blinded for review. The accelerator primarily caters to established organizations, covers the full range of ML technologies, and organizational ML maturity levels.)

2 Use cases and value propositions of machine learning for organizations

2.1 One-off uses of ML within an organization

ML can be used in two types of situations – one-off analysis and repeated use [31]. While the dominant focus in IS and business literature is on the second type, which corresponds to ML-based capability, one-off analysis continues to represent a meaningful share of projects, in which in-house data science teams and consultants engage in. Therefore, our discussion covers both types of ML uses.

One-off analysis utilizing ML, in isolation, does not lend itself to routinization within the context of the STS performing the analysis. Hence, on the level of that STS, there are no new instant capabilities being created. Nor ML becomes a permanent component of the technology underlying STS of the focal capability. The change that is brought by ML typically manifests itself at a lower-level capability and relates to the rearrangement of actors, technology, and structure configuration, or modification of inputs going into that system. More specifically, it is the new insight or knowledge that results from the use of ML and which points to the needed changes in inputs to the STS or STS itself. Such uses of ML potentially bring value to organizations in two ways. First, they can allow organizations to deal with one-off challenges by leveraging new insight or knowledge and existing capabilities. For example, when hurricane Frances was approaching Florida's coastline, Walmart's CIO decided to "start predicting what's going to happen, instead of waiting for it to happen" and mobilized her team to identify which products would be in high demand in the region [32]. Subsequently, Walmart's existing capabilities were used to top up the store shelves with soon-in-demand strawberry Pop-Tarts and beer [32]. This example illustrates that in some cases with new insights generated through one-off ML analysis organizations can leverage existing and unaltered capabilities to resolve a unique problem at hand or benefit from a unique opportunity. Second way, in which organizations can benefit from ML powered one-off analysis is less dramatic but might be even more valuable. New insight or knowledge resulting from the application of ML might be more permanent in nature. For example, one real estate management company relied on an outdated methodology to estimate soil humidity and reimbursed subcontractors for part of their work based on that. (This example was provided on May 24, 2021, by an expert when we were validating the practical relevance of our framework.) When the company received a new ML-based humidity estimation method from consultants, it has turned out that many site types had a dramatically lower humidity than previously expected. This resulted in multimillion-dollar savings on future projects. In this case, ML did not enter into an on-going use by the company, but the insights from one-off analysis improved the overall performance of existing capability and created long-term positive impact on value creation.

P1. One-off use of ML within an organization does not lead to the creation of a new capability.

Returning to the two examples of ML use we discussed above allows us to draw more propositions. While these uses of ML fall into the category of one-off analysis, there is a stark difference between how Walmart generated the new insights compared with the other case. The retail giant relied on an in-house ML capability, while the real estate management company leveraged external ML capabilities. Thus, in both cases ML played a role at a higher-level than operational capabilities. It was used as a tool within the technology element of a higher-level STS. However, Walmart retained that STS within its own organization, while the other company ran one-off projects using external resources. Furthermore, Walmart had all the pieces of the STS needed to carry out the analysis in place, thus demonstrated a routinized process. Hence, we conclude that Walmart had an ML capability. This contrasts with the other company, which not only didn't have the required resources in-house, but also had to carry out non-routine data collection activities to bring the project to fruition. The possession of ML capability within an organization is an important differentiator. Organizations with such capability can not only more rapidly carry out ML initiatives, but also are likely to identify opportunities and deliver on these with higher success rate. Furthermore, multiple executions of one-off ML initiatives might develop or strengthen organizational ML capability, by increasing experience and the level of routinization within that system.

P2. One-off use of ML within an organization may lead to the enhancement of an existing ML capability.

P3. Multiple one-off uses of ML within an organization may lead to the routinization and establishment of ML capability.

2.2 Use of ML capability

While Walmart's use of ML in the example we used related to one-off use, the existing organizational ML capability could also be used when pioneering new or enriching existing operational capabilities [23]. For example, Walmart is developing ML-based capabilities to monitor shelves for product restocking and replenishment needs, as well as to spot problems, such as spills [33]. These operational capabilities are being developed within their Intelligent Retail Lab, which in that context is the higher-level system possessing ML capability. Thus, ML capability can contribute to both one-off uses of ML as well as development of ML-based capabilities. In both cases, the availability of an established and routinized STS underlying ML capability provides an advantage, when compared against organizations without such capability. The case of early collaboration between an external team of researchers and the Atlanta Fire Rescue Department on Firebird serves as a good counterfactual illustration of how the lack of in-house ML capability can undermine ML-based capability establishment. Firebird is a "framework to help municipal fire departments identify and prioritize commercial property fire inspections, using machine learning, geocoding, and information visualization" [34, p. 185]. At the time of writing, Madaio and colleagues concluded that due to poor data sharing practices of the relevant municipal departments, part of the ML development process would need to be redone regularly. Without that the system could not capture changes in the activities and locations of business operating in the commercial properties. Consequently, at the initial phase and due to lack of previously established ML capability, Firebird turned out to be a one-off ML use, which was beneficial, although at that stage did not become an initially envisioned ML-based capability.

P4. Use of an existing ML capability by an organization positively influences the probability of successful outcomes from ML initiatives, including one-off uses of ML and ML-based capability development, in that organization.

However, one-off use of ML by an organization is possible even without having established in-house ML capability. Also, repeated use of ML, thus, an ML-based capability, can be developed without the possession of an ML capability. The most common examples falling within that category are those where ML technology is incorporated into the third-party tools being used within the STS underlying the operational capabilities of an organization. This includes, for instance, prediction of sales leads conversion into opportunities by sales managers using Salesforce Einstein. While this example of tool-like use of ML within the scope of existing operational capabilities has been packaged as a service by cloud vendors, more complex uses of ML, which require, for instance, physical changes in equipment, can also be developed without an in-house ML capability. For example, a German mass producer of electronic sensors and actuators has been developing an ML-based automated visual inspection capability for use by their quality control team by leveraging a collaboration with a university [35].

> *P5. Possession by an organization of an ML-based capability does not require or imply the possession of an ML capability by that organization.*

2.3 Creation of ML-based capabilities

The possession of required, yet disjoint elements of an STS is not sufficient for the establishment of ML-based organizational capability. What is needed beyond these elements is the routinization of their joint activity, to the extent that performance has reached sufficient level of reliability. Such level of routinization is marked by the achievement of stability or balance in the deep structure of the underlying STS. For example, in case of Firebird, due to the one-off nature of data cleaning and joining [34], there was no routinization of the tasks. Thus, the organizational capability to identify and prioritize property fire inspections was not turned into being ML-based. A counter example is that of a global ship brokering company based in Norway, which developed a new ML-based capability to produce oil trade tables – "spreadsheet documents which contain information about activities of certain ships, including timestamps of departures and arrival, destinations, and in which ports they loaded or discharged cargo" [36, p. 6]. The creation of that capability required not only development and integration with existing systems, but also developmental iterations with the maritime activity researchers and redefinition of their role in the process. The establishment and routinization of that ML-based capability took approximately two years.

> *P6. Creation of a new ML-based organizational capability requires not only the presence of suitable actors, social structures, tasks, and technologies, but also establishment of a balanced deep structure linking these elements into a socio-technical system.*

Creation of ML-based organizational capabilities is often rooted in previously existing organizational capabilities. Therefore, ML-based capability creation can often be seen as development or renewal of existing capabilities [12]. Since, by definition, existing capabilities exhibit certain level of balance in their STS's deep structure, introduction of new technological element poses potential threat to that balance. In this section we explore, in the context of first-time ML technology introduction, the relationship between capability performance and stability of its STS.

Returning to the example of ML-based capability of the Norwegian ship brokering company, allows us to elaborate the case of STS being pushed off balance. Since from the initial phase of the project the intention was "to have the algorithm clean, prepare and classify 'raw' AIS data, similarly to what the researcher manually did to generate the tradetables" [36, p. 7], change in the role of maritime researcher was expected, as trade tables' generation was the main responsibility for that job. Yet, at the outset it was not guaranteed that the ML-capability would have superior performance. Thus, unbalancing the STS of an existing capability is not a sufficient condition for performance improvement. This point is also illustrated by the case of a large European company – a member of a global fast-moving consumer goods group with annual revenue of over $50 billion – which aimed at removing "subjectivity and bias from workforce decisions, by drawing on data science, neuroscience, and machine learning" [37, p. 2]. Despite ambitious hopes, the introduction of ML into the trainee recruitment process in Europe resulted in pushing the STS of the underlying capability off balance, while producing disappointing results in terms of improvement of fairness in the selection and recruitment process. Not only did some of the candidates contest the fairness of the process, but also the hiring line managers, the in-house AI team, and the HR managers, who originally spearheaded the project. This imbalance resulted in conflicts between hiring managers, who couldn't hire their preferred candidates, and HR managers defending the ML-based decision rationale.

P7. Unbalancing of an existing capability's STS by introducing to it ML technology element is not a sufficient condition leading to performance improvement of that capability.

While unbalancing of an STS by the introduction of ML is not sufficient by itself to generate capability performance improvement, we posit that it is nonetheless a necessary condition for a performance improvement that is significant. This is because a significant change in the relationship between inputs and outputs of a capability means that the focal STS needs to undergo (or has undergone) a reconfiguration allowing it to exhibit a new range of responses and emergent properties [13]. Such reconfiguration implies not only a substitution of some existing technology with ML, but rather a more encompassing change within the scope of the focal STS. The introduction of ML, in that case, leads to change in one or multiple other elements of the focal STS. Thus, changes in the other elements of the technology, role of actors, social and organizational structure, or underlying tasks are always associated with significant capability performance improvements stemming from the introduction of ML. For example, the case of a Norwegian ship brokering company enriching its tradetable generation capability by introducing ML vividly demonstrates the change in the role of actors and the nature of tasks [36]. After the successful transition to ML-based generation of tradetables the maritime activity researchers became the "teachers and supervisors" of ML algorithm generating the tables used by the rest of the organization. In another study investigating the introduction of ML to Chinese e-commerce giant's fulfillment center – Alibaba's smart warehouse [5] – the traditional areas for manual handling of goods using forklifts and manual labor were replaced by an automatized tridimensional storehouse, where for safety and efficiency reasons people are normally not allowed to enter, position of individual pallets with goods is calculated using an ML-based prediction of the demand for these goods, and robots transport pallets to their destinations. Once orders for goods are received, employees do not need to move around the warehouse to collect goods from a single order, but rather robots do that based on an ML algorithm's probability estimates of various items from multiple real-time ordered being bought together.

P8. Unbalancing of an existing capability's STS by introducing to it ML technology element is a necessary condition for a significant performance improvement of that capability.

2.4 Learning in and improvement of ML-based capability

Having covered the creation of ML-based capability, we shift the focus to subsequent improvements and learning that might take place in such capability. An established ML-based capability must have reached certain level of reliability, has been practiced and routinized, and thus exhibits balance in the underlying STS. Further improvements in the performance of that capability are not guaranteed, despite ML having learning in its name. This notion is evident from even a cursory investigation of ML lifecycle [38], where an ML model deployed into production may go into a new round of learning (re-training), but does not have to. It is thus in the hands of those who develop the ML-based capability to determine whether, how often, and in what form such re-training might take place. In other words, ML model training and inference are two distinct phases in the model lifecycle, and at least one round of training (learning) must take place for an ML to be able to carry out inference in production. For example, a drone capable of object detection and tracking [39] has gone through a training phase and is able to carry out inference using on-board software and hardware. If an organization incorporated such drone into its surveillance or visual inspection capabilities, it could potentially improve performance of these capabilities. However, continuous use of that drone would not by itself result in any changes in the ML algorithm embedded into the drone.

P9. Ongoing use of an ML-based capability by an organization does not require or imply presence of a learning loop, which would improve performance of the ML technology within the STS of that capability.

Thus, organizations developing ML-based capabilities often recognize the need and require keeping ML models up-to-date and, potentially, continuously learning. This is especially the case in the context of high environmental dynamism. Thus, learning feedback loops are often integrated into the overall technology element of the underlying STS. They can take the form of (1) offline maintenance activity or (2) online updating [38]. For example, a European bank periodically retrains their customer service chatbot, which has been developed by an in-house team. (This example was provided by an expert during a workshop, which took place on November 20, 2019.) Such offline maintenance exercise takes place approximately every three months and requires involvement of not only the technical team, but also customer service agents. The retraining targets improvement of ML technology performance related to correctly recognizing customers' intents, as well as updates chatbot's responses, which must correspond to the ever-changing offering and terms of the service. An illustration of an ML-based capability, which has an integrated online learning feedback loop, is the case of Chinese petrochemical plant using digital twin system to control processes of a catalytic cracking unit [40]. Within that system, real-time operational data is not only used by ML to find optimal production settings, but also continuously serves as an input to automatic retraining of the ML model.

While these examples indicate that a feedback learning loop integrated into an established ML-based capability can lead to performance improvement, they do not elucidate the magnitude of changes that are expected or feasible. To complement our discussion with respect to this, we turn to two currently prominent areas of ML development – large transformer-based language models and autonomous driving. OpenAI, an artificial intelligence research and deployment company released in 2020 its third-generation of a large language model called Generative Pre-trained Transformer (GPT-3) [41]. GPT-3 has captured the imagination of media and many practitioners by demonstrating previously unseen performance on multiple tasks. However, plots demonstrating improvements of GPT-3 accuracy for various tasks as a function of number of input parameters consistently show diminishing marginal improvements [41]. Hence, getting ML to perform better gets increasingly harder, as the performance improves. This observation is consistent with the pattern of performance improvements in autonomous driving, where the initial successes were followed by mounting challenges to deal with corner and edge cases [42].

P10. Presence of an integrated learning feedback loop within an established ML-based capability provides, at best, diminishing marginal improvements in the performance of that capability.

Flattening out of capability performance is an expected outcome for all types of capabilities, yet significant performance improvements may be possible even after reaching such plateau [12]. Hence, improvements that would exceed the rate enabled by the integrated learning feedback loop of an ML-based capability might take place in some cases. This, however, requires a new round of development that is substantial and, thus, creates the need to unbalance the established deep structure of the capability's STS. For example, when Airbnb replaced its manual scoring function for property search rankings with a gradient boosted decision tree model, which is a type of ML algorithm, it experienced "one of the largest step improvements in homes bookings in Airbnb's history, with many successful iterations to follow" [43, p. 1927]. However, the algorithm's performance plateaued eventually. This triggered the team responsible for search ranking development to "trying sweeping changes to the system" [43, p. 1927] and introduce a new approach based on deep neural networks. Initially, the team aimed at "keeping everything else invariant and replacing the current model with a neural network" [43, p. 1934], which would retain the existing balance within the underlying STS. This, however, proved to only lower the performance of their search ranking capability. Only by "rethinking the entire system surrounding the model" [43, p. 1934] were they able to gain significant performance improvement.

> *P11. One-off improvements in the performance of an ML-based capability, which are beyond the improvement rate enabled by the integrated learning feedback loop of that capability, may be possible, in which case, to be realized, they require unbalancing of the STS within that capability.*

2.5 Capability automation with ML

In an extreme case, ML, typically in combination with other technologies, can fully automate an existing capability. Such full automation requires a complete encapsulation by technology of four classes of functions, which are (1) information acquisition; (2) information analysis; (3) decision and action selection; and (4) action implementation [44]. Full encapsulation of a capability by technology is possible not only for narrow capabilities, but also in case of socially very complex and core organizational capabilities. Yet even full automation of a capability does not divorce the technology from socio-technical systems constituting an organization. Since STSs are nested structures, encapsulation of a lower-level STS into technology leaves it as an element of technical subsystem of a higher-level STS. Furthermore, the social structure and communication flows within the STS of that higher-level capability may be impacted. Thus, the event of capability encapsulation into ML-based technology may unbalance the STS of the capability being directly above the encapsulated capability. A good illustration of this is provided by the case of a German banking group substituting its in-house capability for small private loan approval and lending term setting with an ML-based technology [45]. Initially, this capability was enriched by incorporating a tool providing "recommendations that loan consultants could change, adapt, or ignore" [45, p. 308]. Later, an enhanced version of the tool, which was ML-based, was implemented as a fully automated solution, which "makes loan approval or denial decisions, determines the terms and conditions of loans, and autonomously alters lending criteria based on customer behavior and current market changes" [45, p. 308]. Thus, the capability for small private loan approval and lending term setting was encapsulated into an ML-based solution, which became an element of technology within the STS of the overarching loan granting capability. This change also brought an upheaval into the social structure within the loan granting capability. The loan consultants, who previously enjoyed relatively high status within the bank because of their experience, required training, certification, and independence in their work, regarded the tool as a threat to their professional role identity and esteem. At the same time, the ML-based tool enabled a new group of employees, such as those working previously at service front desks, receptions, as well as newly hired employees, to promptly assume the role of loan consultants. This equated to a significant professional identity boost for those employees. Furthermore, the use of ML-based loan approval tool erased the need for part of the internal communication flow, which previously served as a document verification step. Thus, the encapsulation of a capability resulted in this case in unbalancing of the STS underlying the loan granting capability of the bank.

P12. Automation of a capability through its encapsulation into an ML-based technology transforms it into an element of the technical subsystem belonging to the STS of a higher-level capability.
P13. Automation of a capability through its encapsulation into an ML-based technology may transform it into an element of the structure within the social subsystem belonging to the STS of a higher-level capability.
P14. Automation of a capability through its encapsulation into an ML-based technology may unbalance the STS of a higher-level capability.

3 Discussion and implications

This paper extends the current debate on organizational capabilities. The topic continues to attract attention of IS and strategic management scholars. Despite this interest, understanding of microfoundations underlying organizational capabilities is still limited. This is especially the case in the context of organizational initiatives, which aim at integrating disruptive digital technologies, such as ML, into their capability portfolio. Evidence from empirical surveys on ML use by organizations show that only one in ten ML initiatives meaningfully contribute to value creation [4]. Thus, practitioners are also facing challenges in this context. With the aim of extending our conceptual understanding of organizational capabilities, this paper takes an STS perspective on the microfoundations of capabilities and discusses the resulting insights relevant to organizational ML initiatives. By conceptualizing organizational capabilities as practiced and routinized transformations of inputs into outputs via underlying socio-technical systems, which are nested in layers and evolve over time, our framework enables a granular insight into the process of digital technology integration into the capability portfolio of an organization. Our contributions arise from the integration of insights from organizational capabilities and STS theory literatures, as well from the derivation of propositions centering on the context of organizational ML initiatives.

With respect to practical implications, our framework and propositions provide several insights, which might be counterintuitive to professionals with limited experience in ML initiatives. For example, use of ML in an organization does not imply that the organization has ML capability or has created any new capability (P1 and P5); there is no free lunch with ML – significant performance improvement of a capability thanks to ML requires significant changes in the STS underlying that capability (P8); ML does not learn by default, once it is put into production (P9); feeding more data into ML leads to, at best, gradually decreasing performance gains (P10); and, full automation of a capability does not completely eliminate the need for people (P12).

3.1 Future research

Future empirical research can build on our work in several ways. First, in the context of ML, future research can test hypotheses drawn from our propositions and identify boundary conditions. Second, our conceptual framework is not limited to ML context, thus, can be used in future studies investigating more broadly the impact of digital technologies on organizational capabilities. Third, by explicitly linking microfoundations to organizational capabilities, the framework allows investigation of technology impact on employees and jobs. Overall, such empirical research will not only advance our understanding of organizational capabilities, but also in the context of ML, will help guide managers and their decisions relating to this transformational technology.

References

[1] H. Benbya, T. H. Davenport, and S. Pachidi, "Special Issue Editorial: Artificial Intelligence in Organizations: Current State and Future Opportunities.," *MIS Quarterly Executive*, vol. 19, no. 4, pp. ix–xxi, Dec. 2020.

[2] E. Brynjolfsson and T. Mitchell, "What can machine learning do? Workforce implications," *Science*, vol. 358, no. 6370, pp. 1530–1534. 2017.

[3] D. Zhang et al., "The AI Index 2021 Annual Report," *arXiv preprint arXiv:2103.06312*, 2021.

[4] S. Ransbotham, S. Khodabandeh, D. Kiron, F. Candelon, M. Chui, and B. LaFountain, "Expanding AI's Impact With Organizational Learning," *MIT Sloan Management Review and Boston Consulting Group*, Oct. 2020.

[5] D. Zhang, L. G. Pee, and L. Cui, "Artificial intelligence in E-commerce fulfillment: A case study of resource orchestration at Alibaba's Smart Warehouse," *International Journal of Information Management*, vol. 57, p. 102304, Apr. 2021.

[6] R. Amit and P. J. H. Schoemaker, "Strategic assets and organizational rent," *Strat. Mgmt. J.*, vol. 14, no. 1, p. 33, Jan. 1993.

[7] C. E. Helfat and S. G. Winter, "Untangling Dynamic and Operational Capabilities: Strategy for the (N)ever-Changing World," *Strat. Mgmt. J.*, vol. 32, no. 11, pp. 1243–1250, Nov. 2011.

[8] J. Iden and B. Bygstad, "Managing Digital Transformation with Sociotechnical Micro-Foundations: A Dynamic Capabilities Approach," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, 2021, p. 6462.

[9] M. H. Jensen, "An Explorative Study of Socio-Technical Information Systems Capabilities," in *39th Information Systems Research Conference in Scandinavia, Ljungskile, Sweden*, 2016, p. 16.

[10] A. S. Bharadwaj, "A resource-based perspective on information technology capability and firm performance: an empirical investigation," *MIS quarterly*, pp. 169–196, 2000.

[11] C. E. Helfat et al., *Dynamic capabilities: Understanding strategic change in organizations*. John Wiley & Sons, 2009.

[12] C. E. Helfat and M. A. Peteraf, "The dynamic resource-based view: capability lifecycles," *Strat. Mgmt. J.*, vol. 24, no. 10, p. 997–1010, Oct. 2003.

[13] K. Lyytinen and M. Newman, "Explaining information systems change: a punctuated socio-technical change model," *European Journal of Information Systems*, vol. 17, no. 6, pp. 589–613, Dec. 2008.

[14] G. Gavetti, "Cognition and Hierarchy: Rethinking the Microfoundations of Capabilities' Development," *Organization Science*, vol. 16, no. 6, pp. 599–617, Dec. 2005.

[15] T. Felin, N. J. Foss, and R. E. Ployhart, "The Microfoundations Movement in Strategy and Organization Theory," *ANNALS*, vol. 9, no. 1, pp. 575–632, Jan. 2015.

[16] M. Govers and P. van Amelsvoort, "A socio-technical perspective on the digital era: The Lowlands view," *European Journal of Workplace Innovation*, vol. 4, no. 2, 2019.

[17] K. Lyytinen, J. V. Nickerson, and J. L. King, "Metahuman systems=humans+machines that learn," *Journal of Information Technology*, 2020.

[18] B. Niehaves, R. Plattfaut, and S. Sarker, "Understanding dynamic IS capabilities for effective process change: a theoretical framework and an empirical application," 2011.

[19] F. Emery, "Characteristics of Socio-Technical Systems," in *The Social Engagement of Social Science, Volume 2*, E. Trist, H. Murray, and B. Trist, Eds. University of Pennsylvania Press, 1993, pp. 157–186.

[20] H. J. Leavitt, "Applied organizational change in industry, structural, technological and humanistic approaches," *Handbook of organizations*, vol. 264, 1965.

[21] J. Barney, "Firm resources and sustained competitive advantage," *Journal of management*, vol. 17, no. 1, pp. 99–120, 1991.

[22] D. J. Teece, "Business models and dynamic capabilities," *Long range planning*, vol. 51, no. 1, pp. 40–49, 2018.

[23] D. G. Sirmon, M. A. Hitt, and R. D. Ireland, "Managing Firm Resources in Dynamic Environments to Create Value: Looking Inside the Black Box," *AMR*, vol. 32, no. 1. 2007.

[24] D. Zhang, S. L. Pan, J. Yu, and W. Liu, "Orchestrating big data analytics capability for sustainability: A study of air pollution management in China," *Information & Management*, p. 103231, Nov. 2019.

[25] T. L. Griffith and D. J. Dougherty, "Beyond socio-technical systems: introduction to the special issue," *Journal of Engineering and Technology Management*, vol. 19, no. 2, pp. 205–216, Jun. 2002.

[26] P. M. Leonardi and S. R. Barley, "Materiality and change: Challenges to building better theory about technology and organizing," *Information and Organization*, vol. 18, no. 3, pp. 159–176, Jan. 2008.

[27] E. L. Trist, *The evolution of socio-technical systems*, vol. 2. Ontario Quality of Working Life Centre Toronto, 1981.

[28] S. Winter, N. Berente, J. Howison, and B. Butler, "Beyond the organizational 'container': Conceptualizing 21st century sociotechnical work," *Information and Organization*, vol. 24, no. 4, pp. 250–269, Oct. 2014.

[29] R. P. Bostrom and J. S. Heinen, "MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes," *MIS Quarterly*, vol. 1, no. 3, p. 17, Sep. 1977.

[30] C. J. G. Gersick, "Revolutionary Change Theories: A Multilevel Exploration of the Punctuated Equilibrium Paradigm," *AMR*, vol. 16, no. 1, pp. 10–36, Jan. 1991.

[31] F. Provost and T. Fawcett, "Data Science and its Relationship to Big Data and Data-Driven Decision Making," *Big Data*, vol. 1, no. 1, pp. 51–59, Mar. 2013.

[32] C. L. Hays, "What Wal-Mart Knows About Customers' Habits," *The New York Times*, Nov. 14, 2004.

[33] A. D'Innocenzio, "Walmart experiments with AI to monitor stores in real time," *The Associated Press*, Apr. 25, 2019.

[34] M. Madaio et al., "Firebird: Predicting fire risk and prioritizing fire inspections in Atlanta," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 185–194.

[35] T. Koppe and J. Schatz, "Cloud-based ML Technologies for Visual Inspection: A Case Study in Manufacturing," in *Proceedings of the 54th Hawaii International Conference on System Sciences*, Jan. 2021, p. 1020.

[36] T. Grønsund and M. Aanestad, "Augmenting the algorithm: Emerging human-in-the-loop work configurations," *The Journal of Strategic Information Systems*, vol. 29, no. 2. 2020.

[37] A. Sergeeva and M. Huysman, "Hiring Algorithms: An Ethnography of Fairness in Practice," p. 10, 2019.

[38] C. Paterson, R. Calinescu, and R. Ashmore, "Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges," *ACM Computing Surveys*, 2021.

[39] S. Han, W. Shen, and Z. Liu, "Deep drone: Object detection and tracking for smart drones on embedded system," 2016.

[40] Q. Min, Y. Lu, Z. Liu, C. Su, and B. Wang, "Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry," *International Journal of Information Management*, vol. 49,, 2019.

[41] T. B. Brown et al., "Language Models are Few-Shot Learners," *arXiv:2005.14165 [cs]*, Jul. 2020.

[42] A. Eule, "Are We There Yet? Autonomous Driving Has a Long Road Ahead.," *Barrons Magazine*, May 03, 2021.

[43] M. Haldar et al., "Applying Deep Learning to Airbnb Search," in *Proceedings of the 25th ACM SIGKDD Int. Conf. on Knowledge Discovery & Data Mining*, Anchorage USA. 2019.

[44] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, vol. 30, no. 3, 2000.

[45] F. Strich, A.-S. Mayer, and M. Fiedler, "What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity," *Journal of the Association for Information Systems*, vol. 22, no. 2, p. 9. 2021.

ARTICLE 4

Blockchain-Based Deployment of Product-Centric Information Systems

Juri Mattila[a, d]
Timo Seppälä[a, d]
Pellervo Valkama[b]
Taneli Hukkinen[c]
Kary Främling[d]
Jan Holmström[d]

Affiliations:

[a] ETLA Economic Research, Finland
[b] MyTownDoesGood, Finland
[c] BitPanda GmbH, Austria
[d] Aalto University, Finland

Original publication:

Computers in Industry, Vol. 125(2) (2021), 103342.

Abstract

Collecting and utilizing product life-cycle data is both difficult and expensive for products that move between different industrial settings at various points of the product life-cycle. Product-centric approaches that present effective solutions in tightly integrated environments have been problematic to deploy across multiple industries and over longer timespans. Addressing deployment costs, incentives, and governance, this paper explores a blockchain-based approach for the deployment of product-centric information systems. Through explorative design science and systematic combining, the deployment of a permissionless blockchain system for collecting product life-cycle data is conceptualized, demonstrated, and evaluated by experts. The purpose of the blockchain-based solution is to manage product data interactions, to maintain an accurate single state of product information, and to provide an economic incentive structure for the provision and the deployment of the solution. The evaluation by knowledgeable researchers and practitioners identifies the aspects limiting blockchain-based deployment of solutions in the current industrial landscape. Combining theory and practice, the paper lays the foundation for a blockchain-based approach to product information management, placing design priority on inter-industrial and self-sustained deployment.

Keywords

Product-centric information management, Blockchain, Inter-industrial deployment, Platform sustainability

1 Introduction

Products in use—especially durable and capital goods—are valuable sources of information in many industrial settings (Aitken, Childerhouse, & Towill, 2003; Anderson & Zeithaml, 1984; Kärkkäinen, Holmström, Främling, & Artto, 2003; Rink & Swan, 1979). However, in settings where products move between systems and industrial settings at different points in the lifecycle, product data is rarely effectively collected and used (Lehtonen, Ala-Risku, & Holmström, 2012). Moreover, a combination of information asymmetries and a lack of incentives may even result in supply chain actors destroying data valuable to one another (Ala-risku, 2009).

The concept of *product-centric information management* (Kärkkäinen, Ala-Risku, & Främling, 2003; Meyer, Främling, & Holmström, 2009; Tang & Qian, 2008) was developed to enable multiple actors to share information on product individuals comprehensively over their lifecycle. While significant improvements have been observed in case studies (Bussmann & Sieverding, 2001; Främling, Holmström, Loukkola, Nyman, & Kaustell, 2013; Hribernik, Rabe, Schumacher, & Thoben, 2006; Lyly-Yrjänäinen, Holmström, Johansson, & Suomala, 2016; Rönkkö, Kärkkäinen, & Holmström, 2007), the deployment of product-centric information management as a sustained solution has been challenging. Deployment challenges include, *e.g.* high initial costs, scalability (Leitão, 2009; Tähtinen, 2018; Trentesaux, 2009), and unresolved conflicts of interest regarding platform control and governance (K. Främling, Harrison, Brusey, & Petrow, 2007). Establishing more integrated platform solutions for product data management has been similarly challenging (Naphade, Banavar, Harrison, Paraszczak, & Morris, 2011).

This conceptual paper explores blockchain-based deployment of a product-centric information system. The focus is on the use of blockchain-based functionality (Buterin, 2013; Hukkinen, Mattila, Smolander, Seppälä, & Goodden, 2019; Nakamoto, 2008; Poon & Buterin, 2017; Wood, 2013), such as protocols, crypto-mining payments, and smart contracts to initiate and sustain product data collection and use. The purpose is to conceptualize and demonstrate a solution, where the design priority is on the incentivization of actors to participate in providing item-level product lifecycle information, and reimbursing their efforts by using blockchain technology. This paper contributes to research on viable inter-industrial deployment (Alam & El Saddik, 2017; Naphade *et al.*, 2011) and self-sustained platforms (Blossey, Eisenhardt, & Hahn, 2019; De Filippi & Loveluck, 2016; Mattila & Seppälä, 2018).

2 Literature review

Storing and maintaining data on each product individual over its entire life cycle is not a trivial undertaking. The high initial investment has been identified as a reason for why integrated product data management systems have not been widely adopted by the industry (Leitão, 2009; Trentesaux, 2009). As an alternative, more loosely coupled peer-to-peer solutions have been proposed to share the burden (Främling, Kubler, & Buda, 2014; Kärkkäinen, Holmström, *et al.*, 2003; Kubler, Främling, & Derigent, 2015). However, while the use of a peer-to-peer approach reduces the investment cost of individual actors, it introduces a variety of new challenges for product centric information management, *e.g.* tracking and coordinating the global state of the system, attracting a critical mass of users, as well as facilitating authentication and trust in a decentralized manner (Petkovic & Jonker, 2007; Trentesaux, 2009).

2.1 Product-centric information and blockchain

In the field of product lifecycle management, earlier efforts towards using a peer-to-peer network have mainly been aimed at increasing the interoperability and openness of product data systems (Kubler *et al.*, 2017; Raggert, 2015). However, obtaining guarantees of the satisfactory performance of peer-to-peer networks has been found difficult; Due to the coordination constraints involved, evaluating the global state of a fully decentralized system—and thus predicting its behaviour—can be highly challenging (Trentesaux, 2009). Over the last decade or so, blockchain technology has provided a potential solution to this issue by enabling a single programmatic state to be maintained in peer-to-peer networks in an entirely decentralized fashion (Buterin, 2013; Hukkinen *et al.*, 2019; Poon & Buterin, 2017; Wood, 2013).

Consequently, in recent research literature, several conceptualizations have been drafted for using blockchain-related systems to improve the transparency and traceability (Azzi, Chamoun, & Sokhn, 2019; Caro, Ali, Vecchio, & Giaffreda, 2018; Cole, Stevenson, & Aitken, 2018; ElMessiry & Elmessiry, 2018; Galvez & Mejuto, 2018; Heber, 2017; Heber & Groll, 2018; H. M. Kim & Laskowski, 2018; Kshetri, 2018; Lu & Xu, 2017; Tian, 2016; Westerkamp, Victor, & Axel, 2018; Wu, Li, King, Miled, & Tazelaar, 2017), the sustainability (Bai & Sarkis, 2020; Kouhizadeh & Sarkis, 2018; Nayak & Dhaigude, 2019; Saberi, Kouhizadeh, Sarkis, & Shen, 2019), the cybersecurity and resilience (Banerjee, Lee, & Choo, 2018; Kshetri, 2017; Min, 2019; Papakostas, Newell, & Hargaden, 2019), and the integration and interoperability (Dai, Zheng, & Zhang, 2019; Gordon & Catalini, 2018; Huang, Wang, Yan, & Fang, 2020; Korpela, Hallikas, & Dahlberg, 2017; Miller, 2018; Repository, 2016; Ruta, Scioscia, Ieva, Capurso, & Sciascio, 2017) of supply chain and product data management structures. Some conceptualizations have also been presented specifically for distributed workflow management with blockchain-based smart contracts (Bahga & Madisetti, 2016; Chen *et al.*, 2017; Evermann & Kim, 2019; Leiding, Memarmoshrefi, & Hogrefe, 2016; Leng, Jiang, Liu, Chen, & Liu, 2017; Yu *et al.*, 2018). Furthermore, other closely resembling themes have been touched upon in many adjacent research streams, *e.g.* focusing on the use of blockchain systems for data governance (Liang *et al.*, 2017; Turk & Klinc, 2017) and ownership management (Karafiloski, 2017; Toyoda, Mathiopoulos, Sasase, & Ohtsuki, 2017; Zhang & Wen, 2017).

Despite the vibrant streams of publications on the issue in recent years, little attention has been paid to the challenge of combining solution deployment and sustainability at the inter-industry level. For example, (Elmessiry, Elmessiry, & Elmessiry, n.d.; Lu *et al.*, 2019; Sternberg, Hofmann, & Roeck, 2020) address the problem of successfully deploying a blockchain architecture for increased transparency and trust in inter-organizational supply chains but do not consider inter-industrial, or system-of-systems, integration. Conversely, (Jiang, Fang, & Wang, 2019; Özyılmaz & Yurdakul, 2019; Tijan, Aksentijevi, & Ivani, 2019) discuss using a blockchain-based architecture for creating an inter-industrial backend for the Internet of Things, but do not address the feasibility of solution deployment. (Katuwal, Pandey, Hennessey, & Lamichhane, 2018), on the other hand, briefly acknowledges the potential suitability of using a blockchain system as an incentivization mechanism to deploy a global health information exchange but does not address the solution sustainability aspect. Respectively, (Rajala, Hakanen, Mattila, Seppälä, & Westerlund, 2018) points out the need for self-reinforcing business models for sustainable systems-of-systems, but does not discuss the feasibility of solution deployment.

While potentially sharing a common manufacturing supply chain, product items do not usually follow one uniform chain of ownership throughout their individual lifecycles. Therefore, an inter-industrial perspective combining both effective deployment and self-sustainability is required in order to establish a prominent product-centric information solution, enabling transformational insight into individual product behaviour across national and industrial boundaries.

2.2 Blockchain systems and smart contracts

Blockchain technology is often described as a combination of information technology elements and methods enabling the creation of decentralized, distributed, and replicated digital ledgers. To this end, the technology employs *e.g.* peer-to-peer networking, public-key cryptography, digital tokens, multi-version concurrency control, and a cryptographically concatenated chain of data blocks used to store database modifications (Nakamoto, 2008).

For this paper, we define blockchain systems strictly as 1) open source and open access technology compositions; 2) comprising a non-hierarchical peer-to-peer networks without single points of failure or control; 3) which maintain consensus over cryptographically concatenated, shared and replicated append-only data structures; 4) according to deterministic self-contained consensus algorithms, void of external inputs such as validation by central authorities or off-chain signalling (Slootweg, 2016). In other words, we make a clear distinction between blockchain systems and the more loosely defined concept of distributed ledgers. A strict delineation of this kind is necessary, as the latter do not exhibit the same kinds of properties essential to solution deployment, as will be discussed later in this paper in Section 4.3.2.

In a computational sense, blockchain systems can be characterized as distributed state machines: peer-to-peer networks capable of maintaining a single programmatic state—or consensus—across the entire network and its shared data, without any single participant having authority over another. By employing Turing-complete programming languages, state-changing programs known as smart contracts can be created, stored and executed in the blockchain network to facilitate diverse distributed workflows (Buterin, 2013; "Ethereum Frontier Guide," n.d.; Hukkinen *et al.*, 2019; Poon & Buterin, 2017; Wood, 2013).

Smart contracts can be described as programmatic containers for tokenized assets. Essentially, they are persistent computer programs which have the ability to autonomously govern assets and to execute transactions. Once assets are deposited into a smart contract's address, they cannot be recuperated until the programming logic of the smart contract permits it. The logic of the smart contract itself is protected by the distributed blockchain network: any unauthorized attempt to tamper with its design is obvious, and easily discarded by other participants (Buterin, 2013; "Ethereum Frontier Guide," n.d.; Hukkinen *et al.*, 2019; Poon & Buterin, 2017; Wood, 2013).

By default, the execution environment of blockchain-based smart contracts lifeless. In order to interact with the smart contract's workflow in a state-changing manner, one must compensate the network on a per-operational basis for providing service. These compensations are also used to allocate request priority and to deter aberrant behaviour, such as requesting infinite computational loops. As each network interaction is bundled with its respective payment in this manner, any state-changing activities, such as database writes, are commonly referred to as 'transactions' in the blockchain vernacular ("Ethereum Frontier Guide," n.d.).

For this paper, we define smart contracts as digital computer programs that: 1) are written in computer code and formulated using programming languages; 2) are stored, executed and enforced by a distributed and replicated blockchain network; 3) can receive, store, and transfer digital assets of value; and 4) can execute with varying outcomes according to their specified internal logic (Lauslahti, Mattila, Hukkinen, & Seppälä, 2018).

2.3 Problem summary

Deploying product-centric information management systems over the product life-cycle is cumbersome, regardless of the technical approach, as all parties involved in the product-life-cycle also need to participate in the information management solution. Attaining a critical mass for a digital platform often requires considerable initial investments. To deploy a solution, the participation of at least one market side must be first subsidized to attract other market sides onto the platform *via* indirect network effects (Armstrong, 2006; Caillaud & Jullien, 2003; Hagiu, 2014; Hagiu & Wright, 2015; Katz & Shapiro, 1994). Consequently, in order to compensate the high-risk venture of establishing a solution in the first place, the pricing models often involve significant economic rent, reducing the appeal of participation (Gawer, 2009; Hagiu, 2014; Tähtinen, 2018).

Thus, understandably, the question of control and ownership of a product-centric information system has been at the centre of attention in research and development (K. Främling *et al.*, 2007). Recently, however, the problem of control and ownership has increasingly become reframed as a broader question of viable inter-industry deployment, especially in the research domain of cyber-physical systems (Alam & El Saddik, 2017; Naphade *et al.*, 2011; Porter & Heppelmann, 2014).

In addition to the problems related to deployment, another set of problems arises from the complexity of dynamic multi-industrial environments. The problem with static workflow designs is that in today's economy, supply chain structures are often complex and prone to reconfigurations (Ali-Yrkkö, Mattila, & Seppälä, 2017; Rajala *et al.*, 2018). While at the industry level, the data integrations and the required reconfigurations may be manageable, at the inter-industrial level the complexity in this regard increases exponentially. Therefore, even if all the parties involved were fully motivated to co-operate to their best ability, product data regarding individual product items could still become fragmented due to the information asymmetries involved.

The third problematic dimension is related to the motivation to preserve the product data workflow. So far, neither centralized nor peer-to-peer-based solutions have been able to provide a satisfactory solution to the problem of adequately incentivizing solution sustainability beyond individual commercial interests. While centralized models have suffered from asymmetrical power structures and single-points of failure, peer-to-peer models so far have lacked proper governance models to foster sufficient network effects for the solution to perpetuate (Ahluwalia, 2016).

3 Methodology

The proposal for an improved design presented in this paper was developed and evaluated by using an explorative design science research approach. Design science is a research method well suited for situations where a practical problem and its solution can effectively be examined through the development of a design artefact, such as a computer program, a system model, or a conceptual practice (Holmström, Ketokivi, & Hameri, 2009; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2008). The design science approach was selected because it enables a rigorous way of designing, building, and evaluating a conceptualization for a product-centric information management system.

The study also incorporates elements of the methodology of systematic combining where an emergent theoretical framework, the empirical fieldwork, practical demonstration, and outcome evaluation are developed in a simultaneous, iterative process (Dubois & Gadde, 2002, 2014). While systematic combining is particularly useful for proposing new approaches and ideas for conceptual research, the main focus of this study is in new practice design. It assumes an integrational approach, providing a cross-disciplinary evaluation of the applicability of blockchain technology to address the challenges of introducing product-centric information management in an inter-industrial setting.

A former case study is also exploited and modified to demonstrate some of the key aspects of the conceptualized design proposal (Eisenhardt, 1989). The demonstration was iteratively developed and contextualized to a relevant product item example and industry setting. The programming of this design artifact draws from the methodologies of computer science (Ayash, 2014).

Through an evaluation procedure, design science enables research objectives to be addressed and problematic areas to be charted and pinpointed at an early phase, without waiting for large-scale implementation. To evaluate the validity of the de-

Table 1. A description of the evaluation interviews

SUBJECT	1st ROUND DURATION	2nd ROUND DURATION	AGE	OCCUPATIONAL TITLE	AFFILIATION	EXPERIENCE IN PRODUCT DATA SYSTEMS (YEARS)	EXPERIENCE IN BLOCKCHAIN TECHNOLOGY (YEARS)
#1	51 min	45 min	39	chief technology officer	industry	11	4
#2	68 min	75 min	54	industrial internet facilitator	academic	25	4
#3	61 min	71 min	34	university lecturer	academic	8	0
#4	61 min	45 min	42	entrepreneur	business	20	4
#5	60 min	61 min	55	program manager	industry	25	0
#6	51 min	58 min	24	doctoral candidate	academic	5	2
#7	56 min	51 min	45	head of digitalization	regulator	0	4

sign proposal, and to provide further in-depth insights into the conceptualization, two rounds of seven qualitative interviews were conducted in a semi-structured manner. The interviews were not intended as a substitute for field testing of the design proposal, but for evaluating the key assumptions and concepts, as well as mapping the critical issues related to the implementability of the design. In other words, the aim was to involve the interviewees in exploring what aspects of the problem situation are important from the interviewee perspective, and how these concerns relate to their view and evaluation of the design proposal. A description of how the evaluation sessions were carried out is presented in Appendix A.

The interviewees were selected in an opportunistic fashion, based on their credentials and expertise, and their heuristically evaluated ability to provide the most valuable insights on the design proposal. The first round of evaluation interviews involved a generic system-level demonstration which was not contextualized to any particular product item or industrial setting. The follow-up interview round involved a more detailed and contextualized iteration of the design proposal with a specified product item, a conceptual data model of the product system architecture (not to be confused with a product data model), and an improved source code artefact with more elaborate incentivization and payment mechanisms. The follow-up interviews also involved a Delphi segment (Dalkey & Helmer, 1963) which allowed the interviewees to comment on the summarized key points from the first round of interviews and to readjust their views. The interview questions around which the interviews were framed is included in Appendix B.

4 Solution proposal and demonstration

4.1 Objectives for a solution

On the basis of the problem summary in Section 2.3, we determine that the main objective for a solution is a design for a product-centric information management system which can be deployed across many industries in terms of costs, coordination, and critical mass, and which can sustain its own existence independently. We postulate that in order to achieve such a design, the system should be able to satisfy the following conditions and specifications: Firstly, the design proposal should be able to a) *enable participation* of all the willing parties. In order to achieve this, the system should feature ahierarchical governance. Secondly, the proposal should be able to b) *prevent data and workflow fragmentation* in a dynamic environment. For this purpose, the system should be based on replicated and distributed architecture. Thirdly, the design proposal should be able to c) *ensure data and platform sustainability* over the complete lifespan of product individuals. For this reason, the system should involve an inherent incentivization mechanism.

4.2 Design principles

We address the research problem and our objectives for an improved design with an approach based on blockchain technology. The motivation for choosing this approach stems from the observation that permissionless open source blockchain systems exhibit a range of properties which conveniently line up with our objectives for a solution. Firstly, due to their ahierarchical governance structure, blockchain systems can be well-suited for enabling participation. Secondly, their blockchain data structure and consensus mechanism can be very effective in maintaining multi-version concurrency control in a decentralized fashion. And lastly, crypto-token-based incentivization mechanisms can be directly incorporated in their participation protocol. Furthermore, the chosen approach comes with a proven track record of several peer-to-peer networks already having been successfully deployed in the described manner in the past (*e.g.* Bitcoin, Ethereum).

In order to accomplish our objectives for a solution, the demonstration of the design proposal needs to show that blockchain systems can be used to involve new parties in the product data system. The demonstration also needs to demonstrate that blockchain systems can be used to include new information as a part of the product-centric information management system. Furthermore, the capability for facilitating adequate incentive structures also needs to be demonstrated.

In this paper, we demonstrate these abilities by employing a smart contract to facilitate a product individual's lifecycle journey. The smart contract was designed for Ethereum, as it represents a suitable deployment environment successfully established in a similar manner as conceptualized in this paper. The other option would have been to establish an entirely new blockchain network as a designated deployment environment for product-centric information management. While perhaps better suited for the actual purpose of the use case, this approach would be difficult to demonstrate in a similar capacity and therefore was not pursued in this paper.

In transitioning from product class data to product-centric information management on individual product items, the number of required transactions can be expected to increase many-fold. Furthermore, as individual product items journey through their individual product lifecycles and paths of ownership, the number of information sources and different data system interactions can also be expected to increase heavily. In order to ensure that the data regarding all the product individuals is provided by all the relevant parties, data provision should be directly rewarded at the level of the participation protocol. For seamless inter-industrial functionality, the system should be constructed so that data exchange can happen spontaneously. In other words, no premeditated *ad hoc* data system integrations should be required between the participants, other than with the blockchain network itself. To this end, the demonstration also illustrates how these incentivization mechanisms can be facilitated by a blockchain-based system design. Furthermore, we also conceptualize, how the provision and the development of the product-centric information system itself can be incentivized by a blockchain-based approach.

4.3 Demonstration of blockchain-based deployment: A loader crane for commercial vehicles

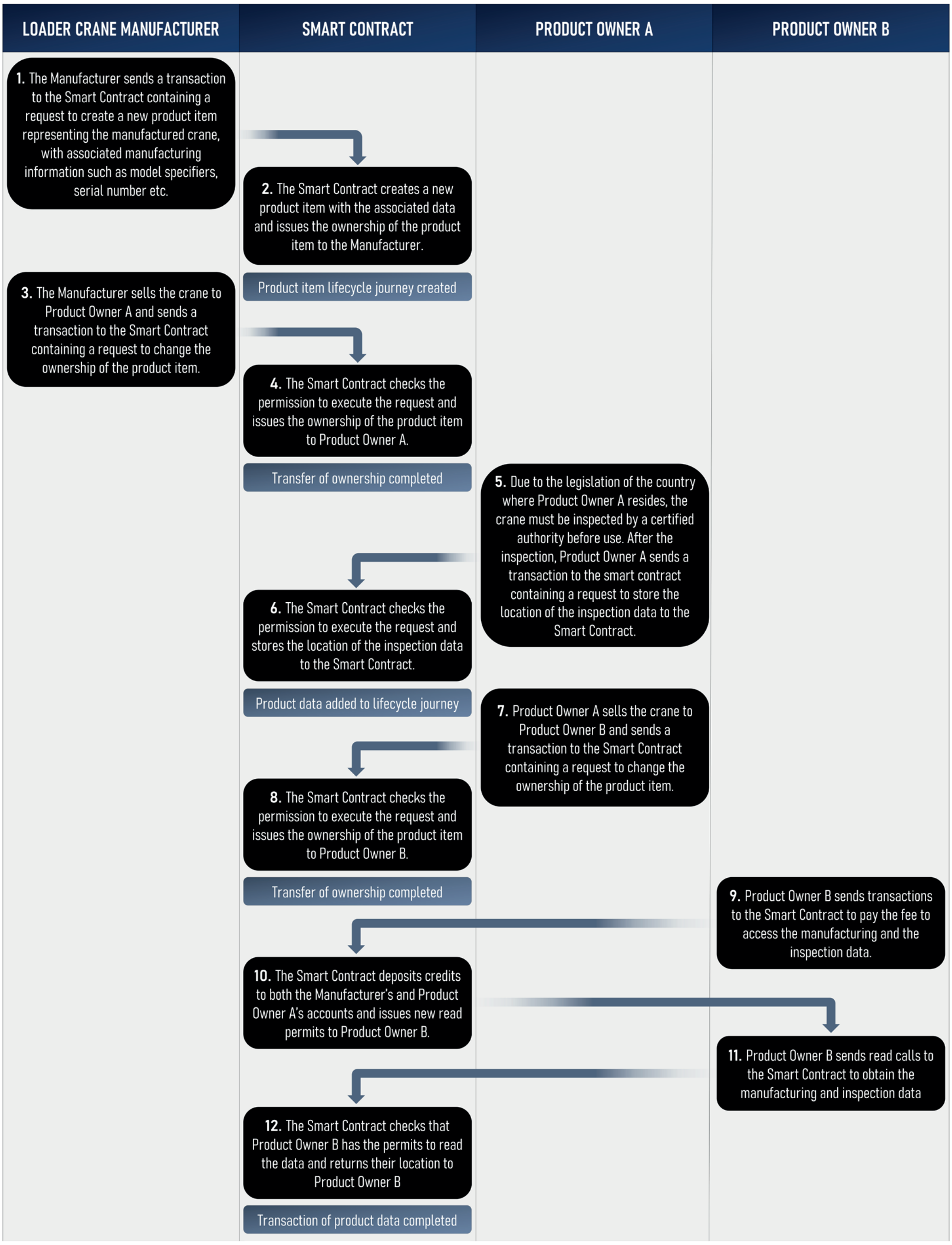
The demonstration of deployment concerns an illustrative product individual, a loader crane for commercial vehicles. These types of loader cranes are manufactured by companies such as Palfinger of Austria, and Hiab of Sweden. The loader crane is typically mounted on a new vehicle before delivery to the customer by the dealer. However, it may also be installed on a vehicle at a later time by the OEM of the loader crane. When the vehicle reaches the end of its life-cycle, the loader crane can be remounted to a different vehicle. This way, the life-cycle of the crane exceeds that of the vehicles to which it is mounted. Over its life-cycle, the loader has many different owners. Furthermore, not only can it be mounted to different vehicles, it can also be repurposed and refurbished by other organizations than the OEM. Product individual data on the loader crane needs to be collected in many countries due to safety regulations.

4.3.1 Participation protocol overview

To demonstrate the conceptualization drafted according to our specified design principles, we present an example protocol of a manufacturer deploying product-centric information management over the product life-cycle of a loader crane (see Figure 1). We demonstrate how the relevant contractual and incentive functionalities in each step are defined in the source code that forms the smart contract in Appendix A. The complete and functional source code for the demonstration can also be found at (Valkama, 2020).

The participation protocol of the demonstration begins with the reception of a new loader crane order by the manufacturer. At this stage, we assume that the smart contract facilitating the workflow for the product life-cycle journey is already deployed in the environment consisting of *e.g.* vehicle manufacturers, loader crane OEMs, truck dealers, trucking firms, and service and maintenance companies. In this conceptualized implementation, after the crane has been manufactured, the manufacturer sends a transaction to the smart contract, requesting that a new product item life cycle journey representing the physical crane is established in the blockchain and its ownership assigned to the manufacturer. In addition, the request contains manufacturing information such as crane model specifiers and a serial number to be stored on the product item (1).

Figure 1 Participation protocol for blockchain-based deployment of product-centric information management over the life-cycle of a loader crane



After this step has been executed by the smart contract (2), the manufacturer can now control the product item in the product data system. As the current owner of the product item, it is possible for the manufacturer to store additional data to the lifecycle journey or query the data already stored without any extra fee.

Upon the sale of the crane to a vehicle manufacturer the crane manufacturer initiates a new transaction in the smart contract in order to transfer the ownership of the product item to the new owner (3). Consequently, the smart contract checks for the permission to perform the request and updates the lifecycle journey accordingly (4).

Over the life-cycle of the loader crane, a multitude of information relevant to different parties is accumulated and can be linked to the smart contract. In the example scenario, once the vehicle manufacturer receives the crane from the loader crane manufacturer, the crane is required to pass an individual inspection performed by a certified authority before it can be installed and used on a vehicle. After the inspection, the vehicle manufacturer sends a transaction to the smart contract in order to store the location pointing to the inspection data (5). Upon receiving the request, the smart contract ensures that the sender of the request is the current owner of the product item and then stores the datum to the smart contract (6).

Once the crane has been mounted onto a vehicle, the vehicle manufacturer delivers the assembly to a truck dealer to fulfil a pre-existing purchase order on the vehicle. Upon the delivery, the vehicle manufacturer sends a transaction to the smart contract in order to transfer the ownership of the product item to the truck dealer (7). The smart contract once again checks for the required permissions and then executes the transfer of the ownership (8).

Before putting the vehicle out for sale, the truck dealer must complete the vehicle registration process and provide documents to the registration authority which prove the vehicle's suitability for its intended use. In order to do this, the truck dealer requires all the relevant information regarding the vehicle's life-cycle journey. To obtain this information, the dealer first sends transactions to the smart contract to pay for the access to the manufacturing and the inspection data from the smart contract (9). Upon receiving the payment transactions, the smart contract deposits credits to the accounts of both the loader crane manufacturer and the vehicle manufacturer for the data they have contributed earlier. Subsequently, the smart contract grants the truck dealer access to the data (10). After the payment transactions have been successfully completed, the truck dealer sends queries to the smart contract to read the relevant data (11). Finally, the smart contract checks that the truck dealer has the valid access and returns the requested data (12). The truck dealer can now proceed with the registration of the vehicle.

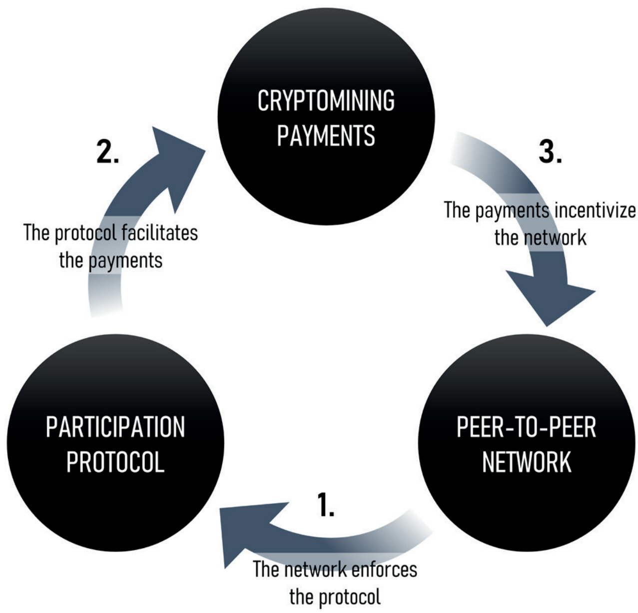
4.3.2 Incentivizing the provision of the product-centric information system

The successful deployment of an inter-industrial product-centric information system, such as the one outlined for the loader cranes, is intricately linked to the concept of *network effects*. In economics, a direct network effect occurs when the value to an agent from using a product, a service or a system depends on the extent of its use by other similar agents. Indirect network effects, in turn, occur when such an increase affects the users of a different product, service or system (Armstrong, 2006; Caillaud & Jullien, 2003; Katz & Shapiro, 1994).

Blockchain-based solutions incorporate a mechanism for a positive feedback loop of indirect network effects to incentivize solution deployment. In essence, the blockchain-based operations described in Appendix A begin by drafting a participation protocol—an elaborate set of rules of engagement to which the participants must adhere in order to be acknowledged by the peer-to-peer network. The actor who initially seeks to create the solution for loader cranes starts the deployment by formulating and publishing the participation protocol. Blockchain systems make use of this participation protocol by inherently embedding financial incentive structures for platform collaboration directly into the protocol itself.

The protocol is open, both allowing new actors to join, as well as the introduction of other types of products than loader cranes. Figure 2 illustrates the positive

Figure 2 The growth-fostering positive feedback loop of network effects in blockchain systems



feedback loop of network effects in blockchain-based deployment. The blockchain system involves a set of rules to which all participants must adhere in order to be acknowledged as members of the network. By contributing computational work, as instructed by the rules of the system, the network enforces a single state of the participation protocol (1). The participation protocol handles each product individual's lifecycle journey and the interactions with it, including the payment transactions for providing product data (2). As each payment also includes a compensation to the network operators for providing service, this incentivization attracts more participants to provide data and to operate the network (3). As the network grows larger, contributing even more computational work (1), the participation protocol grows more robust, making the data and the respective payments in the system more valuable (2). This, again, strengthens the incentives to participate (3), and so on (Athey, Parashkevov, Sarukkai, & Xia, 2016; Athey & Roberts, 2001; Catalini & Gans, 2016; Mattila & Seppälä, 2018).

4.3.3 Incentivizing the provision of product data

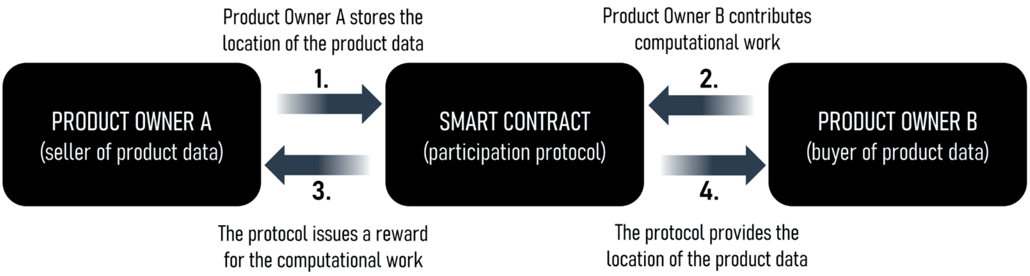
A product datum regarding an individual loader crane can be of very low value to the transacting participants in itself. Therefore, it can be difficult to facilitate the corresponding payments globally in a dynamic environment by any traditional means. Furthermore, in order to maintain the decentralized quality which makes the solution appealing to all parties, the payment processing should also be executed in the same decentralized manner.

While blockchain systems can be used for direct payment processing, they do not scale well in terms of transaction throughput capacity. Therefore, directly facilitating payment transactions through smart contract workflows can quickly become infeasible in large numbers (Hukkinen *et al.*, 2019). Blockchain systems do, however, enable an alternative microtransaction mechanism through the use of crypto-mining payments.

Crypto-mining payments are based on the fact that blockchain systems, require constant inputs of computational work to maintain their single state. Normally, providing this work entitles its contributors to rewards in the form of cryptographic tokens of value in order to incentivize participation. The rewarding is carried out *via* an inflationary tax on the entire network by issuing a small number of new tokens to the recipient of the reward, thus adding tokens into the token supply of the network and depreciating the value of each individual token in the process (Mattila & Seppälä, 2018).

In crypto-mining payments, the cost of the computational work contributed to the network and its respective reward are disentangled from one another to facilitate a payment transaction (see Figure 3). Once the seller has provided the item of

Figure 3 The mechanism of a crypto-mining transaction, as conceptualized in the participation protocol



sale to the smart contract (1), the buyer contributes computational work to maintain the network's concurrency control, expending electricity which effectively constitutes the payment (2). The smart contract then allocates the respective mining reward issued by the network to the seller (3). Finally, the item of sale is delivered to the buyer (4). In essence, in crypto-mining payments, the act of making a payment always simultaneously contributes to the provision of the payment processing platform itself (Pearson, 2018; Rüth, Zimmermann, Wolsing, & Hohlfeld, 2018).

5 Evaluation

5.1 Technical design

The interviewees unanimously considered the loader crane a good product example and an appropriate industrial setting for the conceptualized design proposal. Two of the interviewees commented (#2,3), however, that while the conceptualization seems well-suited for the loader crane—*i.e.* a product of mid-range complexity—in reality product-centric information management must be extended to far simpler products and sub-components than the crane; In such cases, tracking the material and component identities and incentivizing collaboration could become more challenging *via* the conceptualized design, according to the two interviewees. Mostly the interviewees agreed (#1,4,5,6,7), however, that in a full implementation, the participation protocol could be expanded to facilitate the real-world complexity of a product individual's life-cycle.

The final iteration of the participation protocol was considered a sound design and logically coherent by all of the interviewees. One of the interviewees felt (#4), however, that a better possible way of configuring the participation protocol would have been to assign the loader crane product individual with its own unique identity in an equivalent manner to the manufacturer and the owners, and to use the smart contract's workflow only as a transaction link layer for the identities, the data, and the associated payments: "*This, I think, would have been more in line with the current Industry 4.0 digital twin mentality. The added benefit here would be that this participation protocol could guarantee the identities of the agents and product individuals when interacting through this kind of a link layer.*"

As a noteworthy point for further development, one of the interviewees also remarked (#7) on the design proposal's low threshold for extensive field testing: "*One good thing about this conceptualization is that it wouldn't be a huge effort to try this in practice. It's a classic example of a problem that is so complex that it's difficult to anticipate what would happen, so the easiest way to find out would be to simply try it out. And since the concept itself mainly deals with metadata, the risks for the participants would also be quite low.*"

5.2 Enabling participation

In Section 4.1, we postulated that in order to achieve our design objectives, the design proposal should feature ahierarchical governance to enable full participation by all the willing parties. To reflect this design principle, the solution proposal was based on a peer-to-peer blockchain architecture with no centralized authority or any designated individual or group responsible for the solution provision.

The distributed design approach was considered a good and sensible starting point for enabling open participation by all interviewees. Interviewees mostly agreed (#1,4,5,6,7) that successfully establishing an inter-industrial infrastructure at scale will require some new type of an approach. While a caveat offered (#1,6) that starting in the right place does not necessarily mean arriving at a functional solution, the proposed design was generally seen (#1,4,5,6) as a step in the right direction in the design principles. As described by interviewee #4: "*If we think about the loader crane industry, this kind of a systemic approach and the entire platform-building way of thinking is still quite alien to them. However, I think this is the only way to enable vast collaboration between different agents around a single product individual's lifecycle. I don't think any other approach would work at such a high level of scope.*"

The interviewees also largely agreed (#1,4,5,6,7) that the conceptualized open source, open access, and blockchain-based deployment would significantly reduce the costs of solution deployment and lower the barriers of entry into the product data market. The interviewees mostly agreed (#1,4,5,6,7) that the open access design and the role flexibility in solution provision should make participation more inviting, as its less constrictive nature means that participants are free to pursue business opportunities without restrictions by the solution provider. For inter-industrial deployment, this prospect was also considered pivotal (#1,4,6,7) because of the excessive difficulty of any solution provider anticipating all the use cases and business models in which potential participants are interested in an inter-industrial setting. However, arguments were also made (#1,4,6,7) that certain functions could still end up requiring centralized services to be offered on top of the system, involving additional fees for the users; For example, the identities of the users and the product items could turn out difficult to onboard in a completely decentralized fashion.

While the open access to become a provider for the solution architecture was also considered (#2,3,4,6) beneficial for the trustworthiness of the system, one interviewee had (#7) reservations in this regard: "*With this kind of deployment, the network could end up being operated by parties not really involved in the supply chain structures at all. Of course, then you are faced with administrative questions, such as can these parties be trusted and is it really sensible that just literally anyone can start operating the data network. Or do we, after all, want to retain a little bit more control in the hands of those who actually use the data and the system?*"

Some concerns were also raised regarding the scalability of the conceptualized design. These concerns were mainly related to three key points. The first point of concern mentioned (#1,2,5,6) by the interviewees was the possibility of runaway costs due to system inefficiencies as the system is scaled up. This consideration stemmed from the technical properties of the conceptualized solution architecture (*e.g.* the requirement of constant inputs of computational work).

Another point of concern brought up (#1,5,6) regarding scalability had to do with the practical difficulty which often arises in the finer details of scaling up proofs-of-concept and other conceptual solutions. Building conventional IT solutions is a safer practice with a lot more history and experience on avoiding the potential pitfalls. A novel permissionless blockchain-based approach at scale is likely to produce a variety of unforeseeable problems and security issues, such as uncharted attack vectors, which need not have been considered in more traditional approaches.

Lastly, the third scalability-related point of concern mentioned by one interviewee (#2) was the presence of "walled gardens"—the purposeful lack of interoperability maintained by some industry actors as their competitive strategy. Some interviewees felt (#4,6), however, that this kind of a mindset was becoming less common and would be phased out by the market within the next 5–10 years; While customers have not been willing to pay extra for smart product features, market competition is making the smart product approach increasingly a necessity in maintaining a competitive product.

5.3 Preventing data and workflow fragmentation

As our second design objective we stipulated that the system should be based on replicated and distributed architecture in order to prevent data and workflow fragmentation in a dynamic network.

Contemporary solutions to product information management have often involved building case-specific *ad hoc* integrations between the data systems of the vendor and the client. Many of the interviewees expressed (#3,4,5,6) the opinion that due to the difficulty of indexing such *ad hoc* solutions in current configurations, the conceptualized design proposal could help locate the source of product data with greater ease. As explained by interviewee #6: "*When a new system comes along, an integration is built to each pre-existing system. And so the number of APIs absolutely skyrockets, and the system doesn't scale. And at the end of it all, the PLM people are left wondering where the master data is coming from, which systems are integrated with what, and so on. This conceptualization could provide a standard way of transferring the product data between all the various systems.*"

The conceptualized design proposal was purposefully left agnostic in terms of the product data format and meta data standards. The interviewees largely considered (#1,2,3,4,6) this a valid decision, pointing out that specifying a universal standard suitable for the needs of all actors in a cross-industrial context would be exceedingly difficult.

Defining machine-readable formats and relevant meta data standards was, however, considered (#1,2,4,5,6) one of the most important aspects for any shared inter-industrial or even intra-industrial use to be possible. For example, as pointed out by one of the interviewees (#1): "*You want the information fields to have enough flexibility to be able to cover anything, like a potential repurposing of the product, but at the same time, you need enough rigidity to pick up the elements that are important for the loader crane. You need to have the different loader crane manufacturers input similar data in comparable form. That structure is really important.*"

Some the interviewees elaborated (#1,3,4,5,7) that determining such data ontologies was a task best left for the markets and the soft law efforts of each specific industry. As expressed by interviewee #3: "*At the end of the day, everything hinges on what kinds of product data models are demanded by the customers. This way, companies could be forced to switch over to using different kinds of models.*"

In the demonstration's participation protocol, the product data is not stored in the blockchain, as such an approach would hardly be technically feasible. This aspect aroused both positive and negative considerations. The most obvious concern was the fact that the product data still needs to be stored somewhere. While the conceptualization does not describe in detail how the product data could be stored, the interviewees were (#1,4,6,7) open to the exploration of InterPlanetary File System -style solutions. InterPlanetary File System (IPFS) is an open-access peer-to-peer network designed to store data by using content-based addressing. In other words, a given address always points to the same content, thereby preventing data fragmentation within the network[1].

As a positive side, not storing the product data into the blockchain database was seen (#2,4) to enable further access control by each data provider at their end as they see fit. One noteworthy possibility enabled by this aspect, as pointed out (#4) by one of the interviewees, would be the facilitation of product-centric data products. Differing from data-driven applications, such as software solutions using API-based data for analytics, data products are independent, self-adapting entities which combine data inputs with analytical tools and models to produce new outputs of broadly applicable refined data (J. Kim & Bengfort, 2016). Currently, the API-driven solutions utilized in contemporary approaches are insufficient to construct and manage data products effectively. The conceptualized design proposal could offer a way to record and track the product and user identities, ownership relations, and the relevant data ontologies in a more constructive manner.

5.4 Ensuring data and solution sustainability

As the third objective in our design approach, we stated that the system should include an incentivization structure in order to ensure data and solution sustainability over the complete lifespan of product individuals.

One potential problem in this aspect which was pointed out (#1,3,7) is that designing universal incentive structures can be overwhelmingly difficult. For example, if actors were directly compensated for performing transactions of data into product items' life cycle journey, this could lead to the said actors purposefully bloating the system. Similarly, if a generic part of lesser quality is used in maintenance, adding this information to the product data could reduce the resale value of the product. Therefore, the owner may not be inclined to do so, regardless of the incentives embedded in the participation protocol.

While many of the interviewees felt (#2,3,7) that the problems stemming from humans cutting corners cannot be mitigated by incentives embedded in the participation protocol, the resulting market mechanism could alleviate the problem, as explained (#1) by one interviewee: "*If there are 100 fields which should be inputted for the loader crane, is there an incentive to update the fields that are the most popular and have the most valuable use cases? When the system has the incentive mechanism you have conceptualized, I think it will happen organically. When you leave it to a market mechanism, the market will find out which data is more valuable.*"

Another point raised (#2,3,4,6,7) by many of the interviewees regarding the participation protocol was that the system cannot necessarily be perpetuated with internal token incentives alone. Some external motivation for preserving the product data is required outside of the system itself. The interviewees estimated (#1,4,5,6) that the stakeholders in the loader crane's lifecycle would be willing to pay in the order of magnitude of tens to hundreds of euros for relevant data on their product items to be made available upon request, depending on the specific circumstances. This was seen to be motivated by *e.g.* opportunities of increased sales and modernization, regulatory compliance, and reverse logistics at the end of the product lifecycle. Heuristically, the amounts were considered (#1,4,5,6) sufficient to enable the sustained facilitation of the curated workflow, as proposed by the design.

The crypto-mining payments conceptualized in the design proposal provoked a mixed reception. On the one hand, the idea was widely considered intriguing. The notion that every payment transaction also simultaneously contributes to the provision of the underlying payment processing architecture was largely seen (#1,2,4,5,6) as an interesting prospect for fostering positive network effects and producing a positive scaling effect for the deployment of the network. Also, the implications for machine-to-machine payments and the idea that smart devices equipped with some CPU capacity and an internet connection could autonomously pay other devices directly for the curation of their own product data throughout their lifecycles mostly aroused (#2,3,4,5,6) interest.

On the other hand, a majority of the interviewees was concerned (#1,4,6,7) that implementing such a payment model would create an extra layer of unnecessary complexity and token price stability issues, potentially requiring some kind of a middleman to mitigate. Also, in regard to the prospect of M2M payments, it was pointed out (#1,2,3,6) that currently, the vast majority of industrial internet devices in use do not have the required smart capacity to carry out such payments. In the words of interviewee #6: "*Usually the software in products like loader cranes is quite specialized and proprietary, so I imagine adding the capability for crypto-mining payments would be quite a painful endeavour in a larger scale.*"

Due to these considerations, mostly the interviewees largely agreed (#2,3,4,6,7) that while an interesting prospect in its own right, crypto-mining payments would not be feasible as the only possible payment option in the present configuration of industrial systems.

6 Discussion

Several limitations apply which should be acknowledged when interpreting this exploratory study and its findings. Firstly, this study did not explore the integration of the demonstrated design proposal with other IT systems. Secondly, the study did not consider the details of viable product data formats in product-centric information management or the heterogeneity of real-world product data in general. Thirdly, the study did not address the question of how the actor and product identities could be onboarded in a fully decentralized fashion.

The applied semi-structured interview approach is limited in comparison to the more extensive field testing needed for empirical findings and design iterations in accordance with the design science process. The purpose of the loader crane demonstration and its evaluation was not to capture the complexity of a real product lifecycle, however, but to illustrate how a blockchain-based deployment of a product-centric solution could be configured to facilitate the necessary core functionalities for handling the product data, the agent identities, and the incentivization mechanisms required for a full scale implementation. Aiming at a solution that can be deployed across different environments over a long period of time, we seek to contribute to the research on viable inter-industrial deployment (Alam & El Saddik, 2017; Naphade *et al.*, 2011) and self-sustained platforms (Blossey *et al.*, 2019; De Filippi & Loveluck, 2016; Mattila & Seppälä, 2018).

While the use of a blockchain-based system offers a different set of abilities than more conventional approaches, some general problematic aspects regarding its utilization remain which were also not addressed in this paper. For example, while the participation protocol can algorithmically manage the solution provision and the product data workflow, the governance of more strategic development goals remains an open question in the research of blockchain systems (Mattila & Seppälä, 2018). Also, some criticism has also been presented regarding the alleged decentralized nature of blockchain systems in the first place (Walch, 2019).

The proposed approach enables anyone to freely enter the system in any market role and to produce open innovations for all areas and functions of the system. This approach, we anticipate, would create power dynamics where all participants are—not necessarily *de facto* equally powerful—but at least algorithmically equipotent and equally privileged by default. In such a system configuration, no participant would have an obligation to participate in the development, provision, or financing of the system architecture and its auxiliary services, but respectively, no participatory role or function would be off-limits to any participant willing to engage in its provision.

The proposed design presented in this paper extends product data management beyond standard systems. In our proposed design, many such systems are linked in a controlled way, with the product individual as the focal and organizing entity. Even when different actors use their own solutions for product life cycle management information, this information is purposefully collected and distributed between these many systems and actors. Our proposed solution makes it possible to incentivize the collection and distribution of high-quality and high-value product lifecycle information for many different types of product data residing in different systems. This is achieved through a mechanism for different entities to initiate and reward this controlled linking. For example, for a composite product with different modules, the product design and manufacturing information is located in the different PLM systems of the OEMs (*e.g.* Windchill, Teamcenter). The asset and performance data is located in the current and previous owners' operational systems (*e.g.* IBM Maximo, Avantis EAM), and service delivery in the systems of different service providers maintaining and supporting the systems (*e.g.* SAP, Odoo). With the proposed solution, an OEM or a product owner can incentivize other parties to collect and share data on product individuals.

The results of this study suggest that while significant challenges for implementation exist in the current industrial landscape, the applicability of blockchain technology to the problem of product-centric information management has so far been perceived narrowly in academia, largely overlooking its potential significance to sustained inter-industrial deployment. This observation supports the earlier findings of (Blossey *et al.*, 2019) where the authors state that the "[*supply chain*] *applications of blockchain technology mostly focus on efficiency improvements and risk mitigation from a single-firm perspective. – – However, this perspective largely omits the institutional innovation potential of blockchains reorganizing supply chains for collaborative ecosystem-based value creation.*"

The insights provided by this study regarding the incentivized deployment of blockchain solutions for product-centric information management may also help the deployment of similar distributed data sharing solutions intended for other purposes and other sectors of society. The conceptualization delineated in this paper may be especially helpful in cases where the aim is to establish auxiliary services and solutions for business processes that are not core to any of the participants involved. Furthermore, the conceptualized design could also enable an approach where data products on product individuals were manufactured to order, and the curated workflow of the participation protocol served as an index on where the data product could be requested. If successful in its deployment, due to its agnostic data ontology, the system could also be expanded to house a variety of all kinds of data products. Also, the technique could be utilized to manage data in other contexts than product data management, *e.g.* direct from design manufacturing.

7 Conclusion

Our study offers a new network-effect-driven perspective on how inter-industrial data sharing solutions could be established and maintained through a blockchain-based approach, including system development, deployment, and payment processing. In most contemporary design proposals for product-centric information management, the deployment and workflow structures of digital interactions are unilaterally controlled by the service provider who is also providing the underlying technical architecture. By disentangling the solution provision from the control of the data and the workflow, hindrances in the integrational development of inter-industrial digitalization could potentially be alleviated, thus enabling more widespread adoption. Further studies are encouraged for the inter-industrial perspective to product-centric information management, with a design focus on sustained solution deployment.

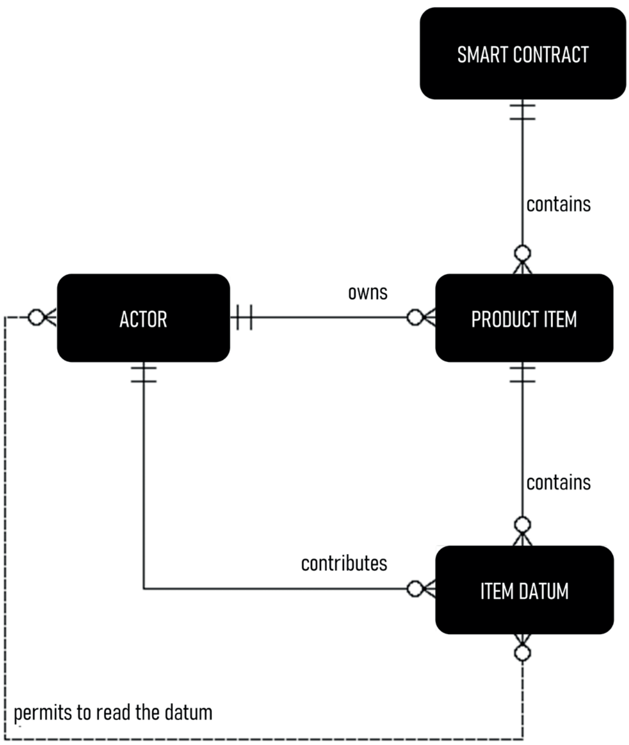
Appendix A: Protocol for blockchain-based deployment

In the following sections, we will present data model, and the different operations that allow the deployment of the loader crane according to the scenario described above. The complete and functional source code for the demonstration can also be found at (Valkama, 2020).

A.1 Product system design

The conceptual data model of the conceptualized system is illustrated in Figure A. The product system contains a collection of product items which are owned by actors such as manufacturers or dealers. The product items each contain a collection of item datums. Consequently, each datum added to a system has an originating actor who is thus considered as the contributor of the datum. Only the contributor of a datum can read the particular datum without cost while all other actors in the sys-

Figure A The conceptual data model of the product system modelled as an Entity-Relationship (ER) diagram



tem are subject to a fee to be able to access it. The actors who have paid the fee are represented in the figure as having the permit to read a datum.

The implementation of the conceptual data model in Solidity, the language used to describe smart contracts in the Ethereum blockchain platform, is shown below:

Product system model (Solidity code)

```
struct ItemDatum {
  address payable contributor;
  mapping(address => uint) permits;
  string datum;
}

struct ProductItem {
  uint itemId;
  address owner;
  mapping(string => ItemDatum) data;
}

uint itemCount;
mapping(uint => ProductItem) private items;
```

The actors in the system are represented simply as Ethereum addresses in the smart contract. This establishes a unique identity to each actor and allows for authentication and access control of the smart contract operations in the Ethereum platform. Furthermore, a simple associative array style data structure of string keys and (datum) values was chosen to represent the product item data. As per the objectives, this imposes minimal restrictions on how to structure and model the product item data, thus enabling different industries to develop their own standards. The requirement of using only textual formats for data also allows for better interoperability across systems and actors. Furthermore, the requirement also discourages polluting the product system with *e.g.* proprietary binary files that are of no use on a larger scale when considering the entire life cycle of a product item and the larger systemic perspective.

The next sections will cover the different operations that are required to implement the semantics of the smart contract, as described in the example scenario. In addition, JavaScript example code of how the smart contract could be called from the client side will be shown.

A.2 Creating a product item life cycle journey

Just as every loader crane in the physical realm goes through a journey of events over its life cycle, respectively, the life cycle of each corresponding product item object in the smart contract can be structured in the same manner. All the product items begin their life cycle journey in the smart contract when a manufacturer sends a transaction to the smart contract, requesting the creation of a new product item with the supplied manufacturing data:

Client side (JavaScript code)

```
createProductItem("4950", {serialNumber: 4950, modelSpecifier: "KPV"});
```

Upon receiving the request sent by the client, the smart contract stores a new product item to the blockchain with the manufacturing data and the sender of the transaction (the manufacturer) as its initial owner. Additionally, the smart contract sends an event, that can be subscribed to by clients, signalling the creation of a new product item:

Smart contract (Solidity code)

```
function createProductItem(
  string memory correlationId
  string memory _manufacturingData
) public returns (uint itemId) {

  uint newItemId = ++itemCount;
  ProductItem memory newProduct = ProductItem({
    itemId: newItemId,
    owner: msg.sender
  });
  items[newItemId] = newProduct;

  setItemDatum(newItemId, "manufacturingData", _manufacturingData);

  emit ProductItemCreated(newItemId, correlationId, msg.sender);

  return newItemId;
}
```

A.3 Transferring the ownership of a product item

When the ownership of a physical loader crane is transferred, the product item in the smart contract must also undergo a transfer of ownership so that the new owner can control the product item. The ownership transfer process is initiated by the current owner by sending a transferral request transaction from the client side to the smart contract, with the product item identifier and the Ethereum address of the new owner as parameters:

Client side (JavaScript code)

```
transferOwnership(4950, "0x485B48DB7e8c65E76178a4C080a7099A5780aA86");
```

Before executing the transfer of the ownership, the smart contract checks that the sender address of the transaction is the same as the address of the owner of the product item. If the sender is not the same as the owner, an error is returned, and the transaction is aborted. After ensuring that the sender is the owner of the product item, the new owner is assigned to the product item and the transaction completes successfully:

Smart contract (Solidity code)

```
modifier onlyOwner(uint productId) {
    require(
        msg.sender == products[productId].owner,
        "Operation permitted only by owner"
    );
    _;
}

function transferOwnership(
    uint _itemId,
    address _newOwner
) public onlyOwner(_itemId) {
    items[_itemId].owner = _newOwner;
}
```

A.4 Assigning new data to a product item

As a loader crane journeys through its individual life cycle, it goes through a unique sequence of transformative events. Respectively, the information contained in the product item must be updated to reflect these changes accordingly. To associate new data to the product item, the owner sends a transaction to the smart contract, using the product item identifier, the key identifying a particular datum, and the datum itself as parameters:

Client side (JavaScript code)

```
setItemDatum(4950, “latestInspection”, {date: “2020-04-21”, result: “ipfs://…”});
```

Upon receiving the request, the smart contract first checks that the sender address of the transaction is the same as the current owner and then updates the product item, associating the datum by its key. Additionally, the address of the sender is stored along the new datum so that the smart contract will later be able to identify the actor who has contributed the particular datum to the system:

Smart contract (Solidity code)

```
function setItemDatum(
  uint_itemId,
  string memory _key,
  string memory _datum
) public onlyOwner(_itemId) {
  items[_itemId.data[_key] = ItemDatum(msg.sender, _datum);
}
```

A.5 Paying to access product item data

If an actor wants to access a particular datum but is not its contributor, the actor must first pay a fee to obtain a right to access the datum. To this end, a transaction is sent from the client side with the product item identifier, the datum key and the payment amount as parameters:

Client side (JavaScript code)

```
payDatumFee(4950, "latestInspection", {value: "1000000000000000"});
```

Upon receiving the payment request, the smart contract first checks that the sender of the transaction is not the contributor of the datum. If the contributor and the sender are the same, the transaction is aborted. Otherwise, the smart contract will deposit the paid fee to the Ethereum address of the contributor and then issue access to the sender while also associating the timestamp of the current blockchain block with the permit:

Smart contract (Solidity code)

```
modifier onlyNotContributor(uint _itemId, string memory _key) {
  require(
    items[_itemId].data[_key].contributor != msg.sender,
    "Only applicable to actors who are not contributors of the datum"
  );
  _;
}

function payDatumFee(
  uint _itemId,
  string memory _key
) public payable onlyNotContributor(_itemId, _key) {
  require(msg.value == 1000000000000000, "Costs 0.001 eth");
  items[_itemId].data[_key].contributor.transfer(msg.value);
  items[_itemId].data[_key].permits[msg.sender] = block.timestamp;
}
```

A.6 Querying product item data

The product item data may be queried at various stages of the product item's life cycle by various different owners. Furthermore, queries can also be made by others actors with access to the smart contract deployment, such as public authorities or third-party integration systems. However, only the original contributor of a particular datum may access it without a cost, whereas other actors must pay a query fee to obtain access. To query data from a product item, a read query is sent from the client side with the product item identifier and the datum identifier as parameters:

Client side (JavaScript code)

```
getItemDatum(4950, “latestInspection”);
```

Upon receiving the query request, the smart contract first checks whether the sender of the transaction is different than the contributor of the datum requested. If the sender and the contributor are the same, the requested datum is returned immediately to the sender. Instead, if the sender and the contributor differ from one another, the smart contract will check whether the sender has access associated with the datum, and in case access has not expired, the datum will be returned:

Smart contract (Solidity code)

```
function getItemDatum(
  uint _itemId,
  string memory _key
) public view returns (string memory datum) {

  if (msg.sender != items[_itemId].data[_key].contributor) {
    uint permitTimestamp = items[_itemId].data[_key].permits[msg.sender];
    require(permitTimestamp + leaseTimeSeconds >= block.timestamp, "No permit");
  }

  return items[_itemId].data[_key].datum;
}
```

Appendix B: Interview guide

TOPIC	KEY QUESTIONS
	Warm-up (1st and 2nd round)
BASIC INFORMATION	▪ Name, age, occupation?
EXPERIENCE	▪ In number of years, how would you describe your experience in: ▪ product data systems? ▪ blockchain technology?
CLARITY	▪ Do you have any questions about the concept?
SENTIMENT	▪ Other initial thoughts about the concept?
	Technical design (2nd round only)
PRODUCT EXAMPLE	▪ How do you feel about the loader crane product item and the industry setting specified for this demonstration?
PARTICIPATION PROTOCOL	▪ What do you think about the technical design of the participation protocol? ▪ Does it make sense to you? ▪ Is there something that jumps out as good or bad? ▪ Is there something that hasn't been considered? ▪ Is there something you would want to change about its design?
FEASIBILITY	▪ How do you see the practical implementability of this design? ▪ How do you feel about its ability to scale and to facilitate the complexity and heterogeneity of real-world product data?
OTHER	▪ Is there anything else you would like to comment about the technical design?
	Conceptualization (1st round: without product & industry context; 2nd round: with said context)
TECHNOLOGY	▪ What potential benefits and problems do you see with the use of a blockchain smart contract to facilitate the workflow of a product life-cycle journey?
GOVERNANCE	▪ The conceptualized PCIM platform has no centralized authority or platform provider. What benefits do you see following from this design principle? What about problems?
DATA FORMAT	▪ The concept does not specify any particular product data format. What are your thoughts on this? Benefits? Problems?
DEPLOYMENT	▪ What do you think about viability of the suggested method of platform deployment through an incentivized open-source participation protocol? ▪ Could you also comment the cross-industrial aspect? ▪ What kinds of problems might the concept solve in establishing a PCIM system? ▪ What kinds of problems might the concept not solve in establishing a PCIM system?

PAYMENTS	▪ What do you think about the suggested method of incentivizing the provision of platform data through crypto-mining payments? ▪ The crypto-mining payment approach would, in principle, enable intelligent product items would be able to pay for the maintenance of their own product data with electricity and CPU power. What are your thoughts on this prospect? What are the benefits and the problems?
LONGEVITY	▪ The concept suggests that due to the incentivization mechanism, the conceptualized PCIM platform could outlive product individuals and even the companies that manufactured them. What benefits and problems do you see with this idea?
VERSATILITY	▪ Due to the open-access nature of blockchain systems, the concept should be able to maintain the product data workflow intact even in the case of dynamic supply chain structures. What benefits do you see to this approach? What about problems?
SHORTCOMINGS	▪ What do you consider the weakest aspect of the concept? ▪ Are there considerations which the concept fails to take into account?
	Delphi (2nd round only)
DELPHI	▪ Do any of these summarized key points in this list jump out to you as something you want to comment? For example, is there something in particular you strongly agree or disagree with?

Endnote

1 For additional information, see <https://docs.ipfs.io/introduction/>. Accessed on 21st of January 2020.

References

Ahluwalia, G. (2016). Device Democracy. The Next Generation IoT Technologies Using the Blockchain. Samsung Developer Connection. Retrieved from https://youtu.be/hwaBM-kQeqc

Aitken, J., Childerhouse, P., & Towill, D. (2003). The impact of product life cycle on supply chain strategy. *International Journal of Production Economics, 85*(1), 127–140.

Ala-risku, T. (2009). *Installed Base Information: Ensuring Customer Value and Profitability after the Sale*. Helsinki University of Technology.

Alam, K. M., & El Saddik, A. (2017). C2PS : A Digital Twin Architecture Reference Model for the Cloud-based Cyber-Physical Systems. *Journal of LaTeX Class Files, 14*(8), 1–13.

Ali-Yrkkö, J., Mattila, J., & Seppälä, T. (2017). *Estonia in Global Value Chains*. Helsinki. Retrieved from https://pub.etla.fi/ETLA-Raportit-Reports-69.p

Anderson, C., & Zeithaml, C. (1984). Stage of the Product Life Cycle, Business Strategy, and Business Performance. *The Academy of Management, 27*(1), 5–24.

Armstrong, M. (2006). Competition in Two-sided Markets. *The RAND Journal of Economics, 37*(3), 668–691.

Athey, S., Parashkevov, I., Sarukkai, C., & Xia, J. (2016). *Bitcoin Pricing, Adoption, and Usage: Theory and Evidence*.

Athey, S., & Roberts, J. (2001). Organizational Design: Decision Rights and Incentive Contracts. *American Economic Review, 91*(2), 200–205.

Ayash, E. M. M. (2014). *Research Methodologies in Computer Science and Information Systems*.

Azzi, R., Chamoun, R. K., & Sokhn, M. (2019). The power of a blockchain-based supply chain. *Computers & Industrial Engineering, 135*, 582–592.

Bahga, A., & Madisetti, V. K. (2016). Blockchain Platform for Industrial Internet of Things. *Journal of Software Engineering and Applications, 9*(10), 533–546.

Bai, C., & Sarkis, J. (2020). A supply chain transparency and sustainability technology appraisal model for blockchain technology. *International Journal of Production Research, 58*(7), 2142–2162.

Banerjee, M., Lee, J., & Choo, K.-K. R. (2018). A blockchain future to Internet of Things security: A position paper. *Digital Communications and Networks, 4*(3), 149–160.

Blossey, G., Eisenhardt, J., & Hahn, G. J. (2019). Blockchain Technology in Supply Chain Management: An Application Perspective. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences* (Vol. 6, pp. 6885–6893).

Bussmann, S., & Sieverding, J. (2001). Holonic control of an engine assembly plant: an industrial evaluation. *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace* (*Cat.*

No.01CH37236), *5*, 3151–3156.

Buterin, B. V. (2013). *Ethereum White Paper – A Next Generation Smart Contract and Decentralized Application Platform*.

Caillaud, B., & Jullien, B. (2003). Chicken & Egg: Competition among Intermediation Service Providers. *The RAND Journal of Economics, 34*(2), 309–328.

Caro, M. P., Ali, M. S., Vecchio, M., & Giaffreda, R. (2018). Blockchain-based Traceability in Agri-Food Supply Chain Management: A Practical Implementation. In *oT Vertical and Topical Summit on Agriculture – Tuscany* (*IOT Tuscany*) (pp. 1–4). Tuscany.

Catalini, C., & Gans, J. (2016). *MIT Sloan School of Management Some Simple Economics of the Blockchain Some Simple Economics of the Blockchain* (No. 5191–16). MIT Sloan Research Paper.

Chen, S., Shi, R., Ren, Z., Yan, J., Shi, Y., & Zhang, J. (2017). A Blockchain-based Supply Chain Quality Management Framework. In: *The 14th IEEE International Conference on e-Business Engineering*. Shanghai, China.

Cole, R., Stevenson, M., & Aitken, J. (2018). Blockchain Technology: Implications for Operations and Supply Chain Management. *Supply Chain Management: An International Journal, 24*(4), 469–483.

Dai, H., Zheng, Z., & Zhang, Y. (2019). Blockchain for Internet of Things: A Survey. *IEEE Internet of Things Journal, 6*(5), 8076–8094.

Dalkey, N., & Helmer, O. (1963). An Experimental Application of the Delphi Method to the Use of Experts. *Management Science, 9*(3), 458–467.

De Filippi, P., & Loveluck, B. (2016). The Invisible Politics of Bitcoin: Governance Crisis of a Decentralised Infractructure. *Internet Policy Review, 5*(3), 1–28.

Dubois, A., & Gadde, L.-E. (2002). Systematic combining: an abductive approach to case research. *Journal of Business Research, 55*(7), 553–560.

Dubois, A., & Gadde, L. (2014). Systematic combining – A decade later. *Journal of Business Research, 67*(6), 1277–1284.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review, 14*(1), 57–74.

Elmessiry, M., & Elmessiry, A. (2018). Blockchain Framework for Textile Supply Chain Management. In: Shiping Chen, H. Wang, & L.-J. Zhang (Eds.), *Blockchain – ICBC 2018* (pp. 213–227). Cham: Springer International Publishing.

Elmessiry, M., Elmessiry, A., & Elmessiry, M. (n.d.). Dual Token Blockchain Economy. The Garment Use Case. In: J. Joshi, S. Nepal, Q. Zhang, & L.-J. Zhang (Eds.), *Blockchain – ICBC 2019. Lecture Notes in Computer Science. Conference paper*. (Vol. 11521, pp. 157–170). Springer International Publishing.

Ethereum Frontier Guide. (n.d.). GitBook. Retrieved from https://legacy.gitbook.com/book/ethereum/frontier-guide

Evermann, J., & Kim, H. M. (2019). Workflow Management on the Blockchain – Implications and Recommendations. *Computing Research Repository*, 1–27.

Retrieved from http://arxiv.org/abs/1904.01004

Främling, K., Harrison, M., Brusey, J., & Petrow, J. (2007). Requirements on unique identifiers for managing product lifecycle information: Comparison of alternative approaches. *International Journal of Computer Integrated Manufacturing, 20*(7), 715–726.

Främling, Kary, Holmström, J., Loukkola, J., Nyman, J., & Kaustell, A. (2013). Sustainable PLM through Intelligent Products. *Engineering Applications of Artificial Intelligence, 26*(2), 789–799.

Främling, Kary, Kubler, S., & Buda, A. (2014). Universal messaging standards for the IoT from a lifecycle management perspective. *IEEE Internet of Things Journal, 1*(4), 319–327.

Galvez, J. F., & Mejuto, J. C. (2018). Future challenges on the use of blockchain for food traceability analysis. *Trends in Analytical Chemistry, 107*, 222–232.

Gawer, A. (2009). *Platforms, Markets and Innovation*. Cheltenham, UK: Edward Elgar Publishing Inc.

Gordon, W. J., & Catalini, C. (2018). Blockchain Technology for Healthcare : Facilitating the Transition. *Computational and Structural Biotechnology Journal, 16*, 224–230.

Hagiu, A. (2014). Strategic Decisions for Multisided Platforms. *MIT Sloan Management Review, 55*(2), 71–80. Retrieved from http://sloanreview.mit.edu/article/strategic-decisions-for-multisided-platforms/

Hagiu, A., & Wright, J. (2015). Multi-sided Platforms. *International Journal of Industrial Organization, 43*, 162–174.

Heber, D. (2017). Towards a Digital Twin: How the Blockchain Can Foster E/E-Traceability in Consideration of Model-Based Systems Engineering. In: *21st International Conference on Engineering Design* (Vol. 3, pp. 321–330).

Heber, D., & Groll, M. (2018). A Meta-Model to Connect Model-based Systems Engineering with Product Data Management by Dint of the Blockchain. In: *International Conference on Intelligent Systems* (pp. 280–287).

Holmström, J., Ketokivi, M., & Hameri, A.-P. (2009). Bridging Practice and Theory: A Design Science Approach. *Decision Sciences*, (February).

Hribernik, K. A., Rabe, L., Schumacher, J., & Thoben, K. (2006). The Product Avatar as a Product-Instance-Centric Information Management Concept. *International Journal of Product Lifecycle Management, 1*(4), 367–379.

Huang, S., Wang, G., Yan, Y., & Fang, X. (2020). Blockchain-based data management for digital twin of product. *Journal of Manufacturing Systems, 54*(January), 361–371.

Hukkinen, T., Mattila, J., Smolander, K., Seppälä, T., & Goodden, T. (2019). Skimping on Gas – Reducing Ethereum Transaction Costs in a Blockchain Electricity Market Application. In: *Proceedings of the 52nd Hawaii International Conference on System Sciences* (Vol. 6, pp. 6875–6884). Retrieved from https://

hdl.handle.net/10125/60123

Jiang, T., Fang, H., & Wang, H. (2019). Blockchain-based Internet of Vehicles: Distributed Network Architecture and Performance Analysis. *IEEE Internet of Things Journal, PP*(6), 3.

Karafiloski, E. (2017). Blockchain Solutions for Big Data Challenges A Literature Review, (July), 6–8.

Kärkkäinen, M., Ala-Risku, T., & Främling, K. (2003). The product centric approach: A solution to supply network information management problems? *Computers in Industry, 52*(2), 147–159.

Kärkkäinen, M., Holmström, J., Främling, K., & Artto, K. (2003). Intelligent products: A step towards a more effective project delivery chain. *Computers in Industry, 50*(2), 141–151.

Katuwal, G. J., Pandey, S., Hennessey, M., & Lamichhane, B. (2018). Applications of Blockchain in Healthcare: Current Landscape & Challenges, 1–17. Retrieved from https://arxiv.org/abs/1812.02776

Katz, M. L., & Shapiro, C. (1994). Systems Competition and Network Effects. *Journal of Economic Perspectives, 8*(2), 93–115.

Kim, H. M., & Laskowski, M. (2018). Toward an ontology-driven blockchain design for supply-chain provenance. *Intelligent Systems in Accounting Finance and Management, 25*(1), 18–27.

Kim, J., & Bengfort, B. (2016). *Data Analytics with Hadoop: An Introduction for Data Scientists*. Sevastopol, CA, USA.: O'Reilly Media. Retrieved from https://www.oreilly.com/library/view/data-analytics-with/9781491913734/ch01.html

Korpela, K., Hallikas, J., & Dahlberg, T. (2017). Digital Supply Chain Transformation toward Blockchain Integration. In: *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 4182–4191).

Kouhizadeh, M., & Sarkis, J. (2018). Blockchain Practices, Potentials, and Perspectives in Greening Supply Chains. *Sustainability, 10*(10), 3652.

Kshetri, N. (2017). Can Blockchain Strengthen the Internet of Things? *IT Professional, 19*(August), 68–72.

Kshetri, N. (2018). International Journal of Information Management 1 Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management, 39*(3), 80–89.

Kubler, S., Främling, K., & Derigent, W. (2015). P2P Data synchronization for product lifecycle management. *Computers in Industry, 66*, 82–98.

Kubler, S., Robert, J., Hefnawy, A., Främling, K., Cherifi, C., & Bouras, A. (2017). Open IoT Ecosystem for Sporting Event Management. *IEEE Access, 5*, 7064–7079.

Lauslahti, K., Mattila, J., Hukkinen, T., & Seppälä, T. (2018). Expanding the Platform: Smart Contracts as Boundary Resources. In: *Springer Translational Systems Sciences series, Volume titled as Collaborative Value Co-creation in the*

Platform Economy (pp. 65–90). Springer Singapore. Retrieved from https://link.springer.com/chapter/10.1007/978-981-10-8956-5_4

Lehtonen, O., Ala-Risku, T., & Holmström, J. (2012). Enhancing field-service delivery: The role of information. *Journal of Quality in Maintenance Engineering, 18*(2), 125–140.

Leiding, B., Memarmoshrefi, P., & Hogrefe, D. (2016). Self-managed and blockchain-based vehicular ad-hoc networks. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct – UbiComp '16*, 137–140.

Leitão, P. (2009). Agent-based distributed manufacturing control: A state-of-the-art survey. *Engineering Applications of Artificial Intelligence, 22*(7), 979–991.

Leng, J., Jiang, P., Liu, J., Chen, Q., & Liu, Q. (2017). Self-organizing of Product-oriented Prosumers Manufacturing Community Driven by the Blockchain Technique. *Journal of Guangdong University of Technology, 34*(5), 1–9.

Liang, X., Shetty, S., Tosh, D., Kamhoua, C., Kwiat, K., & Njilla, L. (2017). ProvChain: A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability. In: *17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing* (pp. 468–477).

Lu, Q., & Xu, X. (2017). Adaptable Blockchain-Based Systems. *IEEE Software, 34*(6), 21–27.

Lu, Q., Xu, X., Liu, Y., Weber, I., Zhu, L., & Zhang, W. (2019). uBaaS: A unified blockchain as a service platform. *Future Generation Computer Systems, 101*, 564–575.

Lyly-Yrjänäinen, J., Holmström, J., Johansson, M. I., & Suomala, P. (2016). Effects of combining product-centric control and direct digital manufacturing: The case of preparing customized hose assembly kits. *Computers in Industry, 82*, 82–94.

Mattila, J., & Seppälä, T. (2018). Distributed Governance in Multi-Sided Platforms – A Conceptual Framework from Case: Bitcoin. *Springer Translational Systems Sciences Series, Volume Titled as Collaborative Value Co-Creation in the Platform Economy*.

Meyer, G. G., Främling, K., & Holmström, J. (2009). Intelligent Products: A survey. *Computers in Industry, 60*(3), 137–148.

Miller, D. (2018). Blockchain and the Internet of Things in the Industrial Sector. *IT Professional, 20*(3), 15–18.

Min, H. (2019). Blockchain technology for enhancing supply chain resilience. *Business Horizons, 62*(1), 35–45.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *Www.Bitcoin.Org*, 9.

Naphade, M., Banavar, G., Harrison, C., Paraszczak, J., & Morris, R. (2011). Smarter Cities and T heir Innovation Challenges. *Computer*, 32–39.

Nayak, G., & Dhaigude, A. S. (2019). A conceptual model of sustainable supply chain management in small and medium enterprises using blockchain technology. *Cogent Economics & Finance, 7*(1), 1–23.

Özyılmaz, K. R., & Yurdakul, A. (2019). Designing a Blockchain-Based IoT With Ethereum, Swarm, and LoRa. *IEEE Consumer Electronics Magazine, 8*(2), 28–34.

Papakostas, N., Newell, A., & Hargaden, V. (2019). A novel paradigm for managing the product development process utilising blockchain technology principles. *CIRP Annals – Manufacturing Technology, 68*(1), 137–140.

Pearson, J. (2018). Mining Cryptocurrency to Pay for Journalism Is Not the Worst Idea I've Ever Heard. *Motherboard*. Retrieved from https://www.vice.com/en_us/article/3k7avk/mining-cryptocurrency-to-pay-for-journalism-salon-coinhive-monero-not-the-worst-idea

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2008). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems, 24*(3), 45–77.

Petkovic, M., & Jonker, W. (2007). *Security, Privacy, and Trust in Modern Data Management. Web Services: Concepts, Architectures and Applications*. Berlin: Springer Berlin Heidelberg.

Poon, J., & Buterin, V. (2017). *Plasma: Scalable Autonomous Smart Contracts*.

Porter, M., & Heppelmann, J. (2014). How Smart, Connected Products are Transforming Competition. *Harvard Business Review*.

Raggert, D. (2015). The Web of Things: Challenges and Opportunities. *Computer, 48*(5), 26–32.

Rajala, R., Hakanen, E., Mattila, J., Seppälä, T., & Westerlund, M. (2018). How Do Intelligent Goods Shape Closed-Loop Systems? *California Management Review, 60*(3), 20–44.

Repository, I. (2016). Blockchain ready manufacturing supply chain using distributed ledger, 0–10.

Rink, D., & Swan, J. (1979). Product Life Cycle Research : A Literature Review. *Journal of Business Research, 7*(3), 219–242.

Rönkkö, M., Kärkkäinen, M., & Holmström, J. (2007). Benefits of an item-centric enterprise-data model in logistics services: A case study. *Computers in Industry, 58*(8–9), 814–822.

Ruta, M., Scioscia, F., Ieva, S., Capurso, G., & Sciascio, E. Di. (2017). Poster Abstract: Supply Chain Object Discovery with Semantic-enhanced Blockchain. In: *15th ACM Conference on Embedded Networked Sensor Systems* (*SenSys'17*) (pp. 1–2). Delft, The Netherlands A.

Rüth, J., Zimmermann, T., Wolsing, K., & Hohlfeld, O. (2018). Digging into Browser-based Crypto Mining. In: *Proceedings of the Internet Measurement Conference 2018* (pp. 70–76). Boston, MA.

Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2019). Blockchain technology and its relationships to sustainable supply chain management, *7543*.

Slootweg, S. (2016). Is my blockchain a blockchain? Retrieved from https://gist.github.com/joepie91/e49d2bdc9dfec4adc9da8a8434fd029b

Sternberg, H. S., Hofmann, E., & Roeck, D. (2020). The Struggle is Real : Insights from a Supply Chain Blockchain. *Journal of Business Logistics*, 1–17.

Tähtinen, T. (2018). *How to Turn IIoT Software Solutions into Platforms: A Case Study of Three Existing Solutions*. Aalto University. Retrieved from http://urn.fi/URN:NBN:fi:aalto-201809034820

Tang, D., & Qian, X. (2008). Product lifecycle management for automotive development focusing on supplier integration. *Computers in Industry, 59*(2–3), 288–295.

Tian, F. (2016). An Agri-food Supply Chain Traceability System for China Based on RFID & Blockchain Technology. *2016 13th International Conference on Service Systems and Service Management* (*ICSSSM*), 1–6.

Tijan, E., Aksentijevi, S., & Ivani, K. (2019). Blockchain Technology Implementation in Logistics. *Sustainability, 11*(4), 1–13.

Toyoda, K., Mathiopoulos, P. T., Sasase, I., & Ohtsuki, T. (2017). A Novel Blockchain-Based Product Ownership Management System (POMS) for Anti-Counterfeits in the Post Supply Chain. *IEEE Access, 5*, 17465–17477.

Trentesaux, D. (2009). Distributed control of production systems. *Engineering Applications of Artificial Intelligence, 22*(7), 971–978.

Turk, Ž., & Klinc, R. (2017). Potentials of Blockchain Technology for Construction Management. *Procedia Engineering, 196*, 638–645.

Valkama, P. (2020). Blockchain Product System. Retrieved from https://github.com/peltchu/blockchain-product-system

Walch, A. (2019). Deconstructing 'Decentralization': Exploring the Core Claim of Crypto Systems. In: C. Brummer (Ed.), *Cryptoassets: Legal, Regulatory, and Monetary Perspectives* (pp. 39–68). New York, NY: Oxford University Press.

Westerkamp, M., Victor, F., & Axel, K. (2018). Blockchain-based Supply Chain Traceability: Token Recipes model Manufacturing Processes. In: *2018 IEEE Confs on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics* (pp. 1595–1602).

Wood, G. (2013). *Ethereum: A Secure Decentralised Generalised Transaction Ledger*.

Wu, H., Li, Z., King, B., Miled, Z. Ben, & Tazelaar, J. (2017). A Distributed Ledger for Supply Chain Physical Distribution Visibility. *Information, 8*(4).

Yu, B., Wright, J., Nepal, S., Zhu, L., Liu, J., & Ranjan, R. (2018). TrustChain: Establishing Trust in the IoT-based Applications Ecosystem Using Blockchain. *IEEE Cloud Computing, 5*(4), 12–23.

Zhang, Y., & Wen, J. (2017). The IoT electric business model: Using blockchain technology for the internet of things. *Peer-to-Peer Networking and Applications, 10*, 983–994.

ARTICLE 5

Digital Protocols as Accounting and Incentivization Mechanisms in Anti-Rival Systems:

Developing a Shareable Non-Fungible Token (sNFT)

Esko Hakanen[a]
Ville Eloranta[a]
Jarno Marttila[b]
Sonja Amadae[a, c, d]

Affiliations:

- [a] Aalto University, Finland
- [b] Streamr Network, Switzerland
- [c] Helsinki University, Finland
- [d] Massachute Institute of Technology, USA

Acknowledgements:

This research received funding from the European Commission's EIC Pathfinder project ATARCA (Accounting Technologies for Anti-rival Coordination and Allocation), Grant No. 964678. We thank the consortium members for their support. We thank Timo Seppälä, Juri Mattila, Tomasz Mucha, and Kimmo Karhu for their helpful suggestions and efforts in compiling and editing the book.

Abstract

Decentralized ledger technologies (DLT), such as blockchains, have been primarily designed to facilitate the exchange of unique, scarce items. This paper presents an alternative decentralization protocol based on anti-rival goods, which gain value in repeated use and are not confined by scarcity. We explain the technical approach behind the concept, referred to as shareable non-fungible tokens (sNFTs), and illustrate our argumentation by presenting a pilot case on supporting the community of Streamr–an open-source, decentralized platform for sharing and streaming data. In addition to introducing this new token standard, we contribute to the discussion on the design of decentralized protocols and the growth of digital commons at large.

Keywords

Blockchain, Decentralized ledger technology, Anti-rival, Protocol, Digital commons

1 Introduction

Following the path laid out by Bitcoin, blockchains are commonly perceived as enablers of digital media of exchange in peer to peer (P2P) networks (Nakamoto, 2008). While decentralized ledger technologies (DLT) have been suggested to facilitate new types of economies (Gencer et al., 2018; Lovett & Thomas, 2021; Swan, 2015), the primary emphasis has been on establishing confidence among peers without a centralized authority through an immutable log of transactions (De Filippi et al., 2020). Accordingly, the attempts to develop the technology have followed along this path, e.g., by suggesting ways to utilize and enhance the smart contracts with different functionalities (Mattila et al., 2021; Rajala et al., 2018) or simply combining on-chain and off-chain transactions to reduce resource consumption in logging the exchanges (Hukkinen et al., 2019).[1]

However, surprisingly few proposals have challenged the inherent nature of economic exchange originating from the trade of physical resources. Digital technologies and infrastructures–including DLTs–are socio-technical systems (Nambisan et al., 2020) that reflect the whole society, its structures, and the economic rationale guiding their design (Mindel et al., 2018; Ostrom, 2005). Accordingly, the prevailing economic institutions, including ownership, money, and banking, have evolved to facilitate the structure of our global economy. Such models have been highly effective in describing markets for goods that are essentially scarce, often produced by tapping into a pool of exploitable, limited resources (like physical ones) (Ostrom, 1990). New, even radical, openings are needed that challenge our presumptions and the prevailing economic mechanisms in the design and development of economies based on purely digital goods.

This paper presents a promising approach to a DLT implementation that neither assumes nor requires artificial scarcity. Our insights are based on the work conducted in the ATARCA project, funded by the European Union's Horizon 2020 research and innovation program. The vision of the ATARCA is to *create new decentralized technology*, "anti-rival tokens," and scientifically founded *proposals for new policies* to enable efficient, decentralized, market-style *trading* and *ecosystems* for *anti-rival goods* to address these concerns.[2]

2 Background

A few key concepts are essential to the proposed vision. In particular, the discussions on anti-rivalry, efficiency, and economic systems and institutions provide the background for the vision. Overall, the vision challenges the orthodox economic assumptions and proposes new thinking to facilitate an anti-rival economy for digital goods.

2.1 Anti-rival goods and systems

For decades, economists and other scholars have differentiated between rival and nonrival goods. The basic principle is that rival goods lose value when consumed, whereas nonrival goods may be used repeatedly, without a loss of value (ATARCA, 2022). In Nobel laureate Elinor Ostrom's terms (2005), the value of rival goods will be subtracted upon use, meaning that their subtractability is positive. In contrast, several indications have been made that many digital or information goods have an "*anti*-rival" nature (Kubiszewski et al., 2010; Olleros, 2018). They differ from rival goods as anti-rival goods *gain* value when used, making their subtractability negative. Thus, the underlying economic principles for anti-rival goods are fundamentally different (ATARCA, 2022).

Contradicting the traditional economic thinking on rival resources, which lose value upon use, anti-rivalry focuses on the repeated and expansive use of resources (Weber, 2004). Following Weber (2004), we call these *anti-rival goods* and the incentive and accounting mechanisms that encourage value creation through anti-rival resource sharing *anti-rival systems*. As laid out in Table 1, anti-rival goods can be divided into "network goods," whose subtractability is negative, typically due to network effects, but that are excludable, and "symbiotic goods," whose subtractability is negative and that are non-excludable (Nikander et al., 2020). Notably, both subtractability and excludability are scales. Also, in many cases, the infrastructure on which the resources are handled affects the anti-rival properties of a good: e.g., if a sharing system has a significant transaction cost, a good loses its anti-rival characteristic (Olleros, 2018).

Of course, there are already several kinds of economic structures that are not based on exchangeability. For example, trust and interpersonal (and interorganizational) relationships can be used to organize anti-rival resources in small-scale communities (Barbrook, 1998; Ghosh, 1998). Large institutions can also set open-access

Table 1 The six types of rival, nonrival, and anti-rival goods

	Subtractability		
Excludability	**Rival**	**Nonrival**	**Anti-rival**
Excludable	Private goods (e.g. coffee)	Club/toll goods (e.g. museum visit)	Network goods (e.g. Fortnite)
Non-excludable	Common-pool goods (e.g. ocean fish)	Public goods (e.g. public beach)	Symbiotic goods (e.g. internet)

Source: Nikander et al. (2020).

policies for example in publicly funded research. Moreover, open-source software development has for decades been successful in facilitating anti-rivalry through collective efforts toward a shared goal (Weber, 2004).

However, the mentioned alternative systems have not been without limitations. The systems have either remained on a small-scale (based on interpersonal trust or an agreement of a limited set of actors), relied on institutional power (public funding or policies), or fitted for only some specific context (like open-source software). While there have been efforts in externalizing these structures for more large-scale and mainstream use, such efforts are predominantly prone to the so-called tragedy of commons (Hardin, 1982): failures of collective action happen when the participating entities use up a common resource for their individual gain, resulting in negative externalities and diminishing returns to everyone due to resource overconsumption (Greco & Floridi, 2004; Ostrom, 1990). Clearly, such alternative economic systems have not comprehensively resolved all of our economic systems' limitations.

2.2 Limitations of current economic systems

Our economic institutions, including ownership, money, and banking, have evolved to serve our global, *rival* economy well and the trade of most *nonrival* goods and services somewhat sufficiently. As more and more goods have transformed into digital format (Yoo et al., 2012), markets have failed and changed (Nikander et al., 2020; Nikander & Elo, 2019), and new legislation and new technology have been introduced.[3] However, neither of these have–so far–attempted to transform the underlying logics of value capture and value extraction (i.e., how the value is divided and distributed among the creator and user, respectively). As a result, these markets continue to fail; goods are distributed in an inefficient manner, and the systems might also contribute to increasing inequality.

Anti-rival goods do not fit traditional markets in which supply and demand depend on inherent scarcity. While it has been long argued that information resources need different strategies than other resources (Shapiro & Varian, 1998), efficient markets are still understood under the conditions of perfect competition; when supply and demand are at equilibrium at a market clearing price. However, for goods that have a very high first fixed cost of production, very low marginal cost, and low secondary fixed costs, existing market mechanisms work poorly (Mueller, 2008).

Consider a simple example of a digital resource: (a piece of) information. Thanks to its digital format (i.e., the bits representing the good), basically any holder of that resource can replicate it infinitely. With modern technology, the cost of producing additional copies of the obtained information or data is essentially zero (Weber, 2004; Yoo et al., 2012). This applies especially to anti-rival goods and is closely connected to the challenges of data markets (Koutroumpis et al., 2020; Nikander & Elo, 2019).

New approaches to anti-rivalry can address the problems of the two identified market equilibria associated with digital goods (Koutroumpis et al., 2020; Nikander et al., 2020): either *data is not produced at all*, or the *data is sold at its copying cost.* In terms of allocative efficiency, it has been commonly considered that consumer preferences are best met when consumers can access their desired digital goods at will, paying only the near-zero copying cost. Previous attempts in this field have often related to IPR (Intellectual Property Rights) laws to prevent unauthorized copying of digital goods (Landes & Posner, 1989); without proper structures, the initial production costs of digital goods cannot be covered, disincentivizing the creation of these goods. Hence, the prevalent mechanisms have relied on creating artificial scarcity, limiting the availability of the goods through legislation or technology, thereby leading 1) to, per se, lesser efficiency due to some parties not receiving a copy of the product and 2) to increased enforcement and technology cost. Past research has provided some conceptual models and anecdotal evidence on resolving such issues (Eloranta et al., 2019; Hakanen et al., 2022), while more work is needed (Nikander et al., 2020).

2.3 DLTs for anti-rival incentivization

DLTs and token systems enabled by DLTs provide fertile ground for experimenting and testing the concept of anti-rival tokens. Digital tools allow experimentation with concepts that may be hard to model, quantify, and measure in the analog world, such as anti-rivalry. As previously discussed, digital resources typically have a high marginal cost of production but a low cost of replicating, copying, or sharing. Thus, digital resources facilitate nonrival or even anti-rival characteristics if they are proliferated and shared openly, e.g., in free and open-source projects (FOSS) (Weber, 2004). However, this type of free and open sharing may not always fit the rivalrous market economy, and the anti-rival and nonrival resources are often converted to rival ones by introducing artificial scarcity (Hakanen et al., 2022), e.g., by adopting DRM (Digital Rights Management) technologies.

We contribute to the discussion of alternative economic institutions by presenting an approach based on anti-rival cryptographic tokens. These tokens exhibit anti-rival characteristics designed to capture (at least partially) the anti-rival value of the underlying system. These tokens can be “shared” in the same way anti-rival goods can be shared at minimal transaction cost. The tokens are used to represent quantified anti-rival value that can be accompanied by a qualitative description. In other words, they may function as a store of value or a unit of account that help us to understand *why* the users find those units valuable. The key difference to various other decentralization initiatives is that the tokens are designed to be shared instead of exchanged.

3 ATARCA project and the Streamr Community pilot

This paper is a conceptual article supported by illustrative evidence and results from the ATARCA project. Next, we shortly introduce one of the project pilot cases[4] and the technological approach behind the experiments.

The ATARCA project addresses the challenge of coordinating collective actions within a global digital economy. The focus is on creating cryptographically protected anti-rival tokens, testing their applicability to governing industrial data markets, and fostering cooperation in community-driven currencies. The project has defined three pilot use cases that explore novel incentive mechanisms to capture anti-rival value in different contexts. This paper will focus solely on the Streamr Community case while introducing the common technological approach behind all three pilots.[5] Streamr is a partner in the ATARCA consortium.

Streamr is an open-source platform that aims to create a global decentralized network for open but secure data transfer. The Streamr community members are connected by a shared social goal: the advancement and sustainability of the Streamr project. This goal requires not only technology development but also the *adoption* of it, i.e., use cases in different contexts that successfully adopt P2P technology developed within the Streamr project. Both code and non-code contributions from the community members are valuable for the project.

The Streamr community's underlying challenges relate to the limitations of information commons (de Rosnay & Stalder, 2020; Greco & Floridi, 2004). Open-source software projects or digital commons can also suffer from the "tragedy of commons" (Greco & Floridi, 2004; Hardin, 1982) – a scenario where the short-term benefits of individuals will decrease the value of the open-source community and eventually decay the whole system. In contrast to physical goods commons, information commons do not suffer from overconsumption; they instead become more sustainable through increased consumption due to network effects (Mindel et al., 2018). However, the tragedy of the information commons refers to the eventual collapse of the network when people only consume and no longer contribute to network maintenance (de Rosnay & Stalder, 2020; Greco & Floridi, 2004). These commonly noted challenges of collective action were addressed in the token and system design.

The leaders of the open-source software community (here, the Streamr team) have an interest in screening and protecting the community from low-quality proposals while fairly acknowledging the providers of high-quality efforts. However, the community leaders cannot truly know the future value of any specific contribution. Nor can the leaders know the true preferences of the community members; the screening process rather represents the vision of the leaders. In other words, there is a risk that the leaders are more likely to screen out and reject contributions that are not aligned with their personal views. Thus, open endorsements from community members can boost decentralization, increase transparency, improve coordination of effort, and enable more efficient allocation of resources within the community.

The main aim of the Streamr community pilot is to study and analyze a new incentivization model for reinforcing anti-rival feedback in the ecosystem that underlies the Streamr P2P platform. The specific interest is in incentivizing development contributions in both non-programming (participating in the discourse, sharing knowledge, etc.) and programming (writing and testing code). This experiment introduces a new type of a token that community members can receive and share with others who have also participated in the platform's development.

4 Technological approach: Token and system design

From the technological perspective, ATARCA is developing institutions and incentive systems that are based on cryptographic tokens. This paper presents a *new cryptographic token type*, titled *Shareable Non-Fungible Token* (*sNFT*),[6] which is a specific variant of the already well-known Non-Fungible Token (NFT). NFTs are cryptographical tokens that are unique (at minimum, the tokens have a unique serial number). The smart contract is defined so that each token is uniquely identifiable and separable from others, making the tokens non-fungible as a result. Also, a more specific variant, a *Shareable, Non-Transferable, Non-Fungible Token* (sntNFT), was developed in the project (ATARCA, 2022).

In the Streamr experiment, anti-rival tokens are bespoke cryptographical tokens defined by a smart contract. The identified tokenized incentives are intended to motivate and coordinate ecosystem stakeholders' activities toward the ecosystems' goals. In addition, the sNFT tokens are used to measure the community members' opinions about the desired path of technology development. The shareability of the tokens is an anti-rival feature that is a new protocol to be implemented for a DLT. These tokenized incentives are distinct to different scenarios.

The Streamr Community pilot features three token types build on top of the sNFT: *Contribution* token, *Like* token, and *Endorsement* token.[7] All of these have different mechanisms on how the token functions, incentives actions, and can be earned and utilized. Tokens developed in the Streamr pilot case have gone through a process of ecosystem design that was facilitated through a series of online workshops using collaborative tools (e.g., Miro boards).[8]

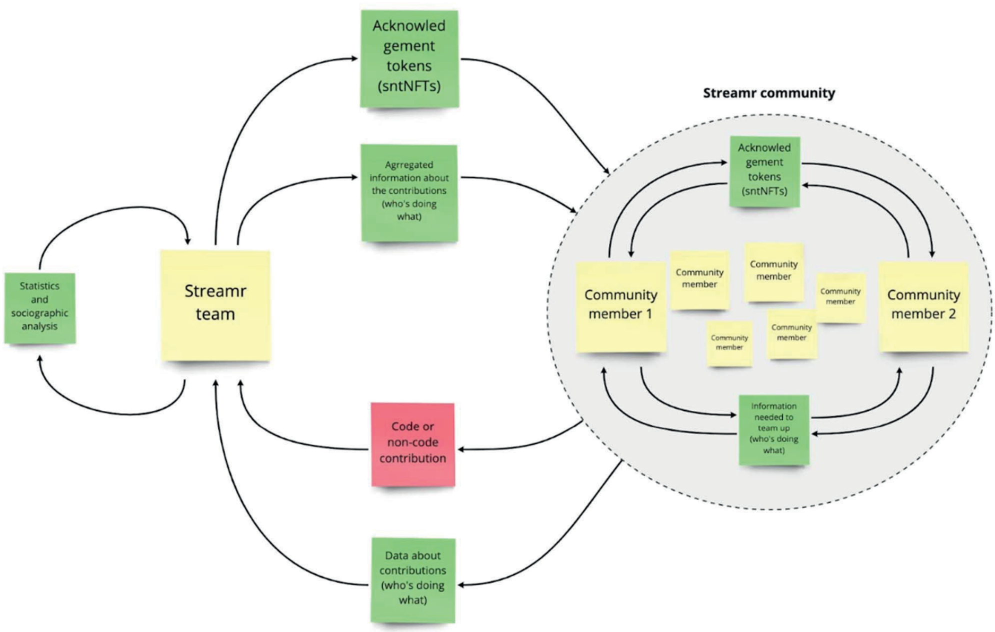
4.1 Mechanism design

Several constraints apply to the Streamr community case mechanism design. Mechanism design has been approached from both the macro and micro levels. On the macro level, we have mapped and reflected macro-level features in relation to sustainability drivers and factors of information commons (Mindel et al., 2018). On the micro level, we have approached the mechanism design and incentive compatibility with the game theoretical approach of a 2x2 game.

The mechanism design aims to actionalize the goals, rules, and incentives of the whole community. In this case, the ecosystem social goal refers to producing valid code- and non-code contributions. What constitutes a valid contribution is specific to the context of the Streamr community pilot experiment use cases. However, all of them should be thought of as an impactful contribution towards the social goal. To improve the sustainability of the digital commons (Mindel et al., 2018), the mechanism design system sought to define a reinforcing loop: the more members the community has, the better the quality of the contributions, the more valuable the community becomes, and the more members will be attracted.

Figure 1 maps and categorizes the value flows between the Streamr team (left-hand side) and Streamr community (right-hand side). Red notes represent flows of

Figure 1 Summary of value flows between the Streamr team and Streamr Community in the Streamr Community pilot. Green items indicate nonrival or anti-rival sharing; red item indicates rival exchange.



rival goods, and green notes flows of non/anti-rival goods. As demonstrated in the figure, self-reinforcing loops emerge between the Streamr team and the Streamr community and inside the Streamr community itself.

The mechanism design seeks to capture these value flows and their positive externalities to maximize the value of the community. The tokens should reflect the identified value flows and actionalize these as tokenized incentives. In the context of anti-rival tokens, incentives are non-monetary, merit-like, and, by definition, 'eternally owned' by their receivers. The shareability function of the sNFT means that Streamr community members who receive a Contribution token are able to share the credit and acknowledge their collaborators essentially by minting a copy of their Contribution token with reference to the original token and appended metadata of the co-contribution, denoting that their contribution has been influenced, affected or contributed to by someone else's contribution. Lastly, a community member can voice their opinion about what contributions they see as valuable by issuing a Like token, or an Endorsement token, given that they already have earned Contribution tokens. These mechanisms help to highlight the merits of a specific contribution (or a Contribution token).

In addition, a linkage between off-chain and on-chain information is utilized to enrich the data stored in the tokens and the DLT. Awarded on-chain tokens are connected to the off-chain metadata to provide further qualitative details of a specific contribution. Metadata is designed to contain information about the type of contribution, e.g., code/non-code, other categorization, receiver (nickname) of the token, a brief natural language description of the contribution itself, and a link to the contribution when possible. When combined with informed consent to release and access metadata, such an approach enables compliance with the general data protection regulation (GDPR), such as the participants' right to be forgotten.

4.2 Token design

DLTs and programmable smart contracts enable us to experiment with new types of digital tokens. Our choice of a DLT platform for the token development has been motivated by its extendability, maturity, and availability of development resources. In the pilot experiment, we have chosen to use NFTs. This choice came from the need to be able to differentiate the tokens from each other and from the need to associate metadata to them when applicable.

The Streamr pilot has different types of NFT tokens in play with different requirements. These tokens have a unique requirement–shareability–which has a different meaning and different implementation depending if the token allows permissioned or open sharing. Shareability is a generic term that can take various forms. For example, one can "share" a digital resource by making a copy of it and by giving it away (share a file), or by agreeing to take turns using one (share a Netflix account with a friend), or one could share a physical resource by giving away a fraction of a whole (share a birthday cake). Thus, the meaning and nature of the sharing depend on its context.

We chose to design and develop tokens on Ethereum Virtual Machine (EVM) compatible smart contracts implemented with Solidity language. EVM is a quasi-turing complete state machine, limited only by the finite number of computational steps available during code execution measured in gas (Antonopoulos & Wood, 2018). Despite the computational limitations, this gives us ample room to explore new types of token implementations.

Current NFT standards do not define nor implement shareability and, hence, have neglected a rational functionality and requirement for any system. The starting point for our technical design and implementation work has been utilizing existing EVM-compatible token standards defining rival tokens, such as ERC-721 NFT "standard implementation" by OpenZeppelin.[9] We focused on removing or adapting elements that impose scarcity and prevent sharing.

We have approached shareability by defining a new Ethereum Improvement Proposal (EIP), EIP-5023.[10] It introduces a new interface that facilitates the creation of shareable NFTs by extending existing NFT contracts with the EIP-5023 sNFT interface (i.e., IERC-5023, Interface of Ethereum Request for Comments). It defines the basic building blocks for sharing – a function method of Share and an event Share. As the meaning of sharing varies between contexts, we believe that the sNFT interface is a valid representation and improvement to current token standards. It leaves the exact implementation of sharing to be handled by its users. At the same time, it enables interoperability between smart contracts as developers can trust that token contracts that use the given interface will behave as defined.

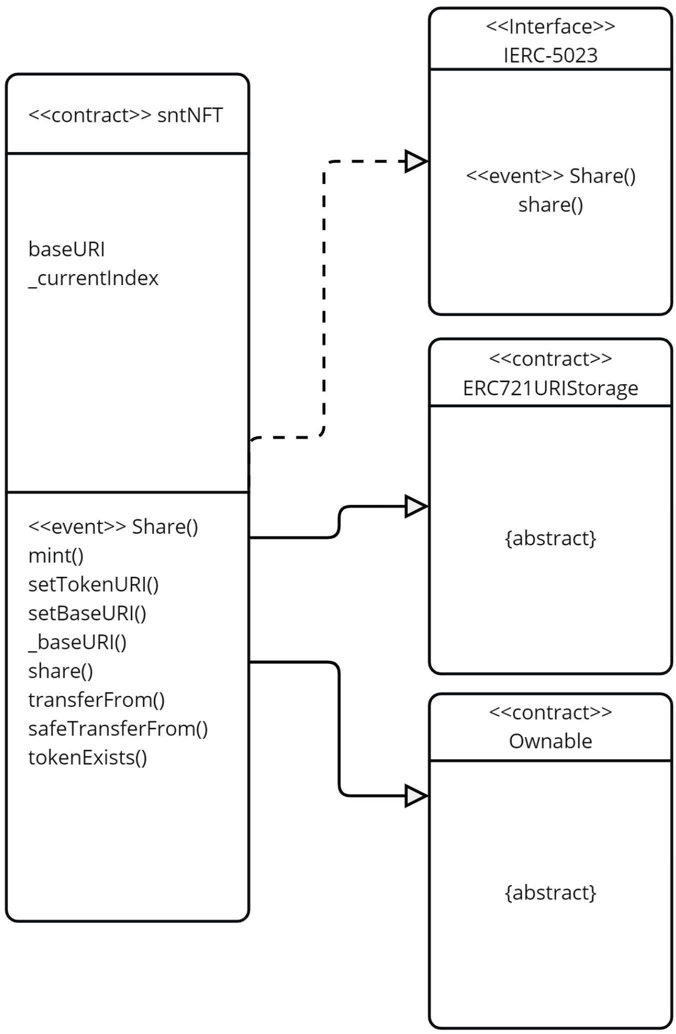
Figure 2 sNFT interface definition

```
/// Note: the ERC-165 identifier for this interface is 0xded6338b

interface IERC5023 is IERC165 {

  /// @dev This emits when a token is shared, reminted and given to another wallet that isn't function caller

  event Share(address indexed from, address indexed to, uint256 indexed tokenId, uint256 derivedFromtokenId);

  /// @dev Shares, remints an existing token, gives a newly minted token a fresh token id, keeps original token at function callers possession and transfers newly minted token to receiver which should be another address than function caller.

  function share(address to, uint256 tokenIdToBeShared) external returns(uint256 newTokenId);

}
```

IERC-5023 share methods expect the function caller to pass two parameters, a wallet or contract address to whom she shares and a token ID to be shared. The function returns a new token ID for the new token minted from the given token and sent to the desired address. A shared event is expected to be emitted during the execution of the share method stating who has shared which token to whom and what is the token ID of the new shared token. Figure 2 summarizes the sNFT interface definition.

In the Streamr Community pilot, the sNFTs have been made *non-transferable* (sntNFT) by overriding transfer-related functions in the contract code. Transfer

Figure 3 The reference implementation of sntNFT



functions are internally usable in the contracts to facilitate token sharing and minting, but they do not allow transferring tokens away from contract users. The sntNFT contract implements the IERC-5023 interface by defining the share "event" and "function" methods. In the reference implementation and in Streamr pilot experiments' contracts, shareability means creating a copy of an existing NFT and giving that copy away at the contract level. This process of copying files the share event of IERC-5023 and a transaction event of IERC-721 contracts conveying that a share has happened and that a shared token has been transferred to a recipient. The events and their associated details are stored in the blockchain's transaction history as log records that can be queried at any time.

Figure 3 presents a UML (Unified Modeling Language) model of the reference implementation of sntNFT, a shareable, non-transferable NFT.[11] The reference implementation builds on top of OpenZeppelin's ERC721URIStorage and Ownable contracts that define NFTs that can have metadata and that contracts can have an owner.[12] The contracts that ERC721URIStorage inherits have been left out of the figure for readability.

4.3 Implementation and governance

The Streamr Community pilot implements and governs three adaptations of the presented token design. Endorsement, Like, and Contribution token contracts implement the IERC-5023 interface and define the sharing functionality in their own contracts. These contracts follow mainly the logic of reference implementation of sntNFTs portrayed in Figure 3. Contribution token contracts access control is set so that only selected members of the Streamr team can mint and transfer Contribution tokens to community members who have successfully contributed to the Streamr community (permissioned sharing). Metadata related to contributions and shared contributions is kept up to date off-chain in a centralized database during the pilot period. Only members who have received contribution tokens can share and re-share their tokens with other community members with the share functionality. Only members who have received Contribution tokens can use Endorsement tokens to support any existing Contribution tokens. However, Like tokens, which reference implementation resembles Endorsement tokens, allow any community member to use Like tokens to support any existing Contribution token (open sharing).

Sharing an Endorsement or a Like token indicates that a person has voiced or shared their opinion with the community by minting "a copy" of a Contribution token to themselves. The Contribution tokens are differentiated between 'original' (minted by the Streamr team) or 'shared' (minted by community members). The contracts for Endorsement and Like tokens query the status of Contribution tokens directly from the Contribution token contract. An Endorsement token has a copy of the Contribution token's metadata appended with a short message from the endorser. Like token does not contain metadata but refers to the Contribution token's metadata when queried. The contract tracks the Likes and Endorsements, and only one Like and Endorsement per wallet address is allowed per each Contribution. Token contracts are built on OpenZeppelin's ERC721Upgradeable token standard. Upgradeability allows the contract owner to change the contract behavior when required. For example, users are able to remove their Likes and Endorsements by burning the tokens they own.

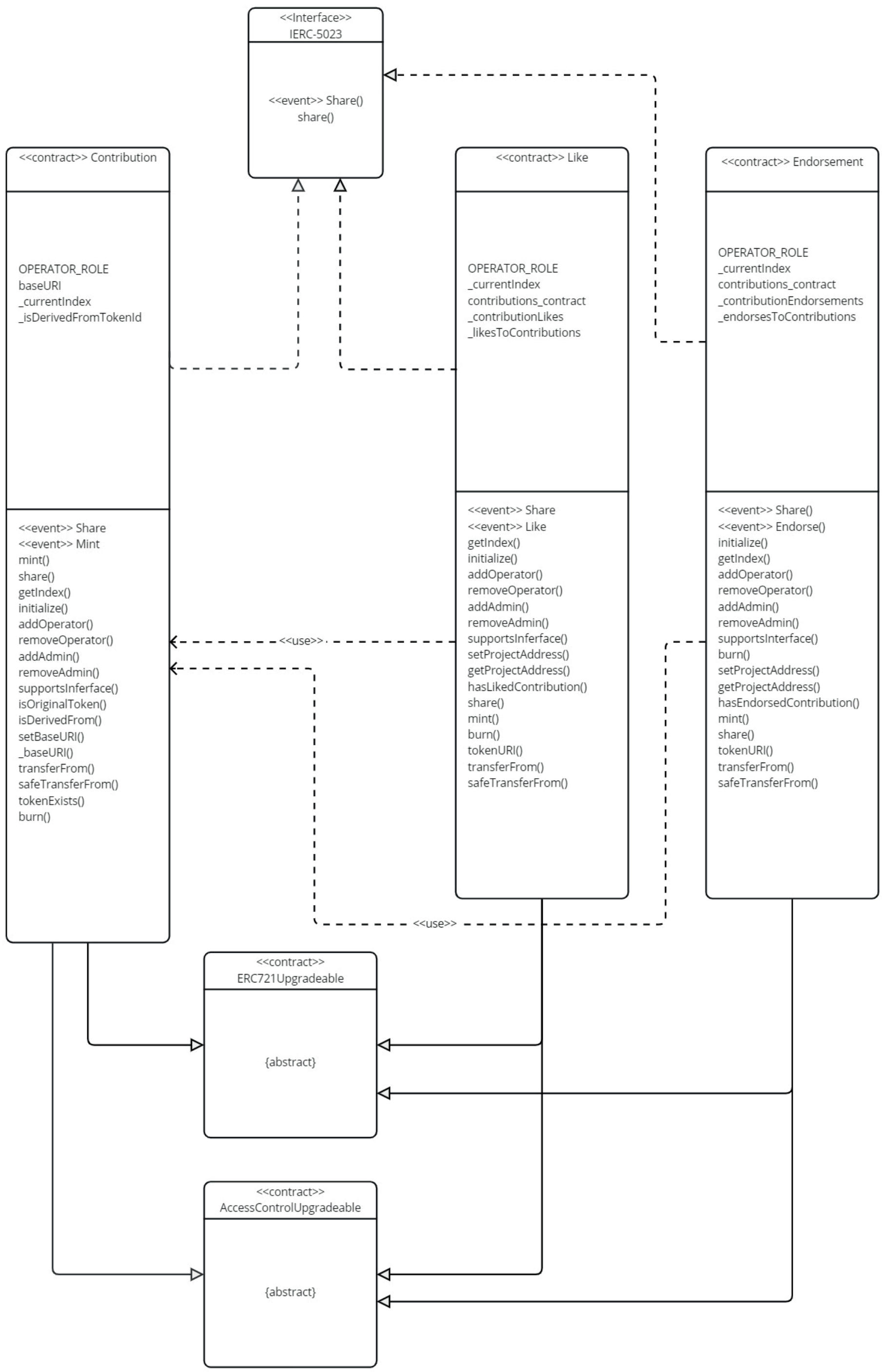
In the Streamr community case, the governance mechanisms are mostly centralized due to the nature of the pilot experiment. The consensus protocol, or the criteria for rewarding the primary Contribution token, is centralized to the Streamr team members responsible of the pilot experiment. The Streamr team establishes the criteria for rewarding a token, and each contribution is evaluated against the criteria. Any conflicts arising in the token system are resolved centrally.

There are different management and governance mechanisms underlying the sNFT tokens. In general, only the owners of the relevant token contracts–selected members of the Streamr team and the research personnel–are allowed to mint and transfer tokens to appraise member contributions. Once Contribution tokens are minted and transferred to their recipients, they cannot be exchanged or transferred away. However, the receiver of a Contribution token can share and transfer it to new owners. Anyone can mint a Like token to themselves as long as the corresponding contribution token continues to exist. Endorsement tokens can only be minted if the minter has an existing Contribution token on her wallet. Like and Endorsement tokens are always linked to a contribution token, thereby maintaining the connection to the original contribution and keeping a record of a growing network of community preferences. Figure 4 shows a UML representation of Contribution, Like, and Endorsement token contracts.

5 Implications

The aim of this paper is to showcase the potential of designing anti-rival systems. The ATARCA project has addressed the issues of open market valuation and the structural disparities in the digital goods and data markets. We believe that such work is needed, as it addresses the root causes related to the market failures of data economy (Nikander & Elo, 2019), poorly working or nonexistent markets for industrial data (Koutroumpis et al., 2020), and many existing data markets reducing to effectively near-zero price (Nikander et al., 2020). In this paper, we focused on how can the crypto-economic mechanisms be used to incentivize the production of anti-rival goods. We have illustrated this work through the process of appraising code and non-code contributions in the Streamr community.

Figure 4 UML model of Contribution tokens, Like tokens and Endorsement tokens contracts



5.1 Token valuation

None of the tokens in the Streamr Community pilot hold any direct monetary value. The main purpose of the tokens is to document the activities and inputs performed by the community members. Hence, they visualize the process leading to an outcome, while enabling a coherent history of previous and linked contributions toward a certain target. Moreover, they provide insights on the views, hopes, and preferences of the community members.

The described system utilizes different instances of the sNFT protocol (Contribution, Like, and Endorsement token) to appraise the work and activities conducted within the system. The transferability of these tokens has been disabled by choice in the design of the incentive mechanism (hence the notion of non-transferability, or sntNFT). This prevents a monetary exchange of these tokens and speculation towards a financial reward, which is found to be a common issue in cryptocurrencies (de Rosnay & Stalder, 2020; Kher et al., 2021). Hence, the approach differs from the predominant view on the design of decentralized protocols that has focused on the tokenization of value in an effort to produce scarce accounting units to be exchanged (Hakanen et al., 2022).

Nevertheless, contribution and endorsement tokens are expected to hold indirect value and capture at least some of the positive externalities arising in the community. Tokens are expected to derive value from the functionality of the Streamr project and from the interaction and information sharing within the community. Further indirect value can be achieved if these kinds of tokens are later used in other domains outside the Streamr ecosystem. Possible use cases include the acknowledgment of open-source community contributions or the creation of meritocratic governance mechanisms in other decentralized open-source projects.

5.2 Research implications

In this paper, we reflected on the current and evolving understanding of the potential of using crypto-economic mechanisms for incentivizing the production of non-/anti-rival goods, especially in ways that omit the need for artificial scarcity. Moreover, we illustrated how digital tools and infrastructures align the creation and sharing of value with anti-rival and nonrival goods. We modeled their impact on alternative incentive mechanisms while creating new types of crypto-economic tokens to capture (some of) the value of network externalities in digital communities (cf. Karhu et al., 2021).

The consortium has sought to reconsider the foundational structures and institutions of our economic systems, many of which are based on concepts that predate the modern era–such as accounting, ownership, private property, money, and banking.

These concepts still shape the contemporary approaches to our economic models, with the implication that the notions of ownership and exchange are often considered an inseparable components in all economic systems, including cryptocurrencies. However, such tendencies have implicated that new approaches were required to facilitate a global economy for digital goods.

Digital information goods deviate from prevalent economic models because they are inherently nonrival (maintaining their value when copied) (Mueller, 2008; Olleros, 2018; Shapiro & Varian, 1998). They are goods with a very high fixed cost of production for the first unit but a very low marginal cost and low secondary fixed costs for the secondary (replicated) units. Moreover, many digital goods and infrastructures have anti-rival characteristics (increasing their value with shared use) (Olleros, 2018; Weber, 2004). For example, the value of an item, such as a piece of software, often increases as more people use the software (Weber, 2004). Thus, the existing market mechanisms work poorly in describing the transaction of digital information goods. New mechanisms are needed to create proper incentive structures to cover the initial production costs of digital goods for more sustainable and efficient digital economies.

In addition, this work highlights a novel avenue for advancing work on collective action and decentralized communities. The technological protocols presented here provide concrete mechanisms to document the work, for instance, in networks or ecosystems without formal hierarchical structures (Autio et al., 2018; Eloranta et al., 2019). More broadly, our work provides an interesting tangent to exploring independent and autonomous agents motivated by a system-level goal, also known as “meta-organizations” (Gawer, 2014; Gulati et al., 2012).

5.3 Managerial implications

We see that the sNFT token and its practical use cases have the potential to be analogous to the manner in which Bitcoin implementation (Nakamoto, 2008) allowed a broad instantiation of blockchains and cryptocurrencies (Swan, 2015). A notable difference is that, while the value of Bitcoin is based on and confined to an artificial scarcity, the value of the sNFTs will be based on visualizing the underlying human relations, efforts, and the value of different interactions. The value of sNFT tokens reflects how relationships and contributions are developed over time through repeated interactions, benefiting all members and various aspects of the community (Barbrook, 1998; de Rosnay & Stalder, 2020; Hakanen et al., 2022; Weber, 2004). Thus, sNFTs can serve as a metric of value, a medium of sharing, and even a store of credit.

We believe that the crypto-economic mechanisms illustrated with the Streamr Community pilot use case are applicable and generalizable to other Web3 communities. We expect that the technology can facilitate an industry-wide contribution of acknowledging positive contributions beyond the scope of this pilot while addressing (some) of the issues in digital commons (Greco & Floridi, 2004) across the FOSS industry (Weber, 2004).

From a managerial perspective, the monetization of digital goods commonly relies on controlling access rights. In many cases, such policies diminish the benefits and value potential of virtually zero copying costs associated with digital resources (Olleros, 2018; Weber, 2004). Yet, if the data access were completely free, creators of these information resources would have limited incentives to invest in creating and providing the good in the first place (Mueller, 2008; Nikander & Elo, 2019; Shapiro & Varian, 1998). Shareable or anti-rival goods and network externalities likely remain outside the traditional market transactions due to limitations in accounting or rewarding for the generation of anti-rival values. The development of anti-rival tokens and a new distributed ledger accounting system enables one to measure, record, and appreciate the anti-rival value and positive externalities.

6 Limitations and further research

This paper is an early attempt to contribute to the design and modeling of digital protocols supporting anti-rivalry, with potentially important implications for the literature on economic institutions. However, more work is needed to provide a deeper understanding of the economics of digital goods (Autio et al., 2018; de Rosnay & Stalder, 2020), especially at the infrastructural level (Mindel et al., 2018; Olleros, 2018). Herein, we agree with the calls for research on allocative inefficiencies, new types of quantified value, and new institutionalisable means of shared and collaborative governance (Koutroumpis et al., 2020; Lovett & Thomas, 2021; Nambisan et al., 2020; Nikander et al., 2020). We also call for further research on increasing and capturing of positive externalities enabled by the circulation of anti-rivalrous community currencies.

Endnotes

1 Several initiatives have been proposed, see, for instance, “Monoplasma: A simple way to broadcast money to millions of people: https://medium.com/streamrblog/monoplasma-revenue-share-dapps-off-chain-6cb7ee8b42fa” or “Bitcoin Smart Contract 2.0: Trustless contracting by combining on-chain and off-chain transactions:” https://xiaohuiliu.medium.com/bitcoin-smart-contract-2-0-d1e044abed5a

2 ATARCA stands for Accounting Technologies for Anti-Rival Coordination and Allocation (EU H2020 Grant No. 964678), see https://atarca.eu for more details.

3 For instance, consider US Digital Millennium Copyright Act: https://en.wikipedia.org/wiki/Digital_Millennium_Copyright_Act or Digital Rights Management (DRM): https://en.wikipedia.org/wiki/Digital_rights_management

4 ATARCA pilots are referred to as: Barcelona Green Shops; Streamr Community Case; and Food Futures. See “Use Cases” at https://atarca.eu/ for more details.

5 For detailed descriptions and rationale behind the ATARCA pilot use cases, please refer to public project deliverables D1.1 and D2.1 at: https://atarca.eu/

6 ATARCA consortium’s “sNFT” Ethereum Improvement Proposal was made public on Apr 15, 2022, and accepted on Jan 3, 2023, immortalizing it as part of the Ethereum project. The full description can be found at https://eips.ethereum.org/EIPS/eip-5023.

7 For a more thorough description, please visit: https://blog.streamr.network/streamr-awards-are-here-contribute-and-earn-unique-snfts/.

8 We utilized the anti-rival business design toolkit in this work. See: https://github.com/ATARCA/Anti-Rival-Business-Design-Toolkit/.

9 OpenZeppelin: The standard for secure blockchain applications, see: https://github.com/OpenZeppelin/openzeppelin-contracts

10 Reference implementation available on Github at: https://github.com/ethereum/EIPs/blob/master/EIPS/eip-5023.md

11 Reference implementation available on Github at: https://github.com/ethereum/EIPs/blob/master/EIPS/eip-5023.md

12 Details for ERC721URIStorage: https://github.com/OpenZeppelin/openzeppelin-contracts/blob/master/contracts/token/ERC721/extensions/ERC721URIStorage.sol and Ownable: https://github.com/OpenZeppelin/openzeppelin-contracts/blob/master/contracts/access/Ownable.sol

References

Antonopoulos, A. M., & Wood, G. (2018). *Mastering Ethereum*. O'Reilly Media.

ATARCA. (2022). *Report on Crypto-economic Mechanisms for Anti-rival Goods – Project Deliverable 2.1 (D2.1)*.

Autio, E., Nambisan, S., Thomas, L. D. W., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, *12*(1), 72–95. https://doi.org/10.1002/sej.1266

Barbrook, R. (1998). The hi-tech gift economy. *First Monday*, *3*(12).

De Filippi, P., Mannan, M., & Reijers, W. (2020). Blockchain as a confidence machine: The problem of trust & challenges of governance. *Technology in Society*, *62*, 101284. https://doi.org/10.1016/j.techsoc.2020.101284

de Rosnay, M. D., & Stalder, F. (2020). Digital commons. *Internet Policy Review*, *9*(4), 1–22. https://doi.org/10.14763/2020.4.1530

Eloranta, V., Hakanen, E., Töytäri, P., & Turunen, T. (2019). Aligning multilateral value creation and value capture in ecosystem-level business models. *Academy of Management Proceedings*.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, *43*, 1239–1249. https://doi.org/10.1016/j.respol.2014.03.006

Gencer, A. E., Basu, S., Eyal, I., van Renesse, R., & Sirer, E. G. (2018). Decentralization in bitcoin and ethereum networks. *Financial Cryptography and Data Security*. https://doi.org/http://dx.doi.org/10.1016/j.jpain.2010.01.228

Ghosh, R. A. (1998). Cooking pot markets: An economic model for the trade in free goods and services on the internet. *First Monday*, *3*(3).

Greco, G. M., & Floridi, L. (2004). The tragedy of the digital commons. *Ethics and Information Technology*, *6*(2), 73–81. https://doi.org/10.1007/s10676-004-2895-2

Gulati, R., Puranam, P., & Tushman, M. L. (2012). Meta-organization design: Rethinking design in interorganizational and community contexts. *Strategic Management Journal*, *33*, 571–586. https://doi.org/10.1002/smj

Hakanen, E., Eloranta, V., & Shaw, C. (2022). Forming digital commons: The role of signaling and sociology of translation in decentralized systems. *Academy of Management Proceedings 2022*, 11067.

Hardin, R. (1982). *Collective Action*. Resources for the Future.

Hukkinen, T., Mattila, J., Smolander, K., Seppala, T., & Goodden, T. (2019). Skimping on gas – reducing Ethereum transaction costs in a blockchain electricity market application. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, *6*, 6875–6884. https://doi.org/10.24251/hicss.2019.823

Karhu, K., Heiskala, M., Ritala, P., & Thomas, L. D. W. (2021, October). Platform externalities: Beyond the n in network effects. *Druid21*.

Kher, R., Terjesen, S., & Liu, C. (2021). Blockchain, Bitcoin, and ICOs: A review and research agenda. *Small Business Economics*, *56*, 1699–1720.

Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020). Markets for data. *Industrial and Corporate Change*, *29*(3), 645–660. https://doi.org/10.1093/ICC/DTAA002

Kubiszewski, I., Farley, J., & Costanza, R. (2010). The production and allocation of information as a good that is enhanced with increased use. *Ecological Economics*, *69*(6), 1344–1354. https://doi.org/10.1016/j.ecolecon.2010.02.002

Landes, W. M., & Posner, R. A. (1989). An economic analysis of copyright law. *The Journal of Legal Studies*, *29*(2), 325–363.

Lovett, M., & Thomas, L. (2021). A fork in the road: Perspectives on sustainability and decentralised governance in digital institutions. *First Monday*, *26*(11).

Mattila, J., Seppälä, T., Valkama, P., Hukkinen, T., Främling, K., & Holmström, J. (2021). Blockchain-based deployment of product-centric information systems. *Computers in Industry*, *125*. https://doi.org/10.1016/j.compind.2020.103342

Mindel, V., Mathiassen, L., & Rai, A. (2018). The sustainability of polycentric information commons. *MIS Quarterly*, *42*(2), 607–631. https://doi.org/10.25300/MISQ/2018/14015

Mueller, M. (2008). Info-communism? Ownership and freedom in the digital economy. *First Monday*, *13*(4).

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. www.bitcoin.org

Nambisan, S., Lyytinen, K., & Yoo, Y. (2020). Digital innovation: Towards a transdisciplinary perspective. In S. Nambisan, K. Lyytinen, & Y. Yoo (Eds.), *Handbook of Digital Innovation* (pp. 2–12). Edward Elgar Publishing Limited. https://doi.org/10.4337/9781788119986.00008

Nikander, P., & Elo, T. (2019). Will the data markets necessarily fail? A position paper. *30th European Conference of the International Telecommunications Society (ITS): "Towards a Connected and Automated Society,"* 1–9.

Nikander, P., Eloranta, V., Karhu, K., & Hiekkanen, K. (2020). Digitalisation, anti-rival compensation and governance: Need for experiments. *Nordic Workshop on Digital Foundations of Business, Operations, and Strategy*, 1–6.

Olleros, F. X. (2018). Antirival goods, network effects and the sharing economy. *First Monday*, *23*(2).

Ostrom, E. (1990). *Governing the Commons* (J. E. Alt & D. C. North, Eds.; 1st ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511807763

Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton University Press. https://doi.org/10.1007/s11127-007-9157-x

Rajala, R., Hakanen, E., Seppälä, T., Mattila, J., & Westerlund, M. (2018). How do intelligent goods shape closed-loop systems? *California Management Review*, *60*(3), 20–44. https://doi.org/10.1177/0008125618759685

Shapiro, C., & Varian, H. (1998). *Information Rules: A Strategic Guide to the Network Economy*. Harvard Business School Press.

Swan, M. (2015). *Blockchain–Blueprint for a New Economy* (1st ed.). O'Reilly Media. https://doi.org/10.1109/CANDAR.2017.50

Weber, S. (2004). The Success of Open Source. In *The Success of Open Source*. Harvard University Press. https://doi.org/10.2307/j.ctv26071g2

Yoo, Y., Boland, R. J. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, *23*(5), 1398–1408. https://doi.org/10.1287/orsc.1120.0771

ARTICLE 6

The Little Engines That Could:

Game Industry Platforms and the New Drivers of Digitalization

Juri Mattila[a]
Timo Seppälä[a, b]
Kaisa Salakka[c]

Affiliations:

[a] ETLA Economic Research, Finland
[b] Aalto University, Finland
[c] Unity Software Inc., Finland

Original publication:

Mattila Juri, Seppälä Timo and Salakka Kaisa (18.11.2021). "The Little Engines That Could – Game Industry Platforms and the New Drivers of Digitalization". ETLA Brief No 101.

Abstract

In a recent trend in digitalization, many platform incumbents have steered their focus towards creating collectively shared persistent virtual frameworks known as 'metaverses'. Due to the emergence of digital platforms in the game industry over the last decade, the industry is now challenging the digital platform incumbents in metaverse development. Will the development unlock new data-driven markets, how will the landscape of digital platforms be reconfigured, and what are the strategic and policy implications for Finland and the European Union?

Keywords

Game engine, Virtual reality, Metaverse, Platform, Platform Business Group (PBG) Strategy, System of systems, Digitalization

Digital platforms – A gamechanger for the game industry

In the past decade, discussion has been vibrant regarding a new development in digitalization, a so-called 'data economy'. In the discussion on the digital platform economy, however, it has long since been recognized that, at its core, digitalization has more to do with *interactions* than data itself. While data can certainly be valuable, mostly its value derives from enabling more productive interactions between parties, or better-informed decisions regarding those interactions, in one form or another. In this respect, few others have harnessed data to facilitate interactions as prominently as digital platforms in the platform economy (Still et al., 2017).

In the past decade or so, the game industry has undergone a significant transformation in how games are played, developed, and distributed due to the onset of digital platforms. During this time, distribution platforms, such as Google Play, Apple Store, and Steam have opened up an entirely new array of game industry markets. By offering significantly larger developer revenue shares than the former industry standards before platforms, and by enabling access to vastly larger target audiences, the platform giants have enabled smaller game studios to become more empowered in game content creation.

At the same time, platformization has also started taking hold of the game industry in other layers of the technology stack. With creations such as Quake Engine by id Software, Unreal Engine by Epic Games, and RenderWare by Criterion Software, game houses started developing game engines independently from the game content already in the late 1990s. In this transformation, content creation—such as graphics, storyline and characters, and the game-specific rules and objectives—was separated from building the basic game infrastructure—such as the game physics, collision detection, graphics rendering, and networking. By licensing these infrastructural frameworks, or *game engines*, to other game studios as the foundation for new games, some studios were able to tap into an additional lucrative business-to-business revenue stream, while providing other studios with more versatility in game design, lower development costs in development and lower barriers of entry into the game industry markets.

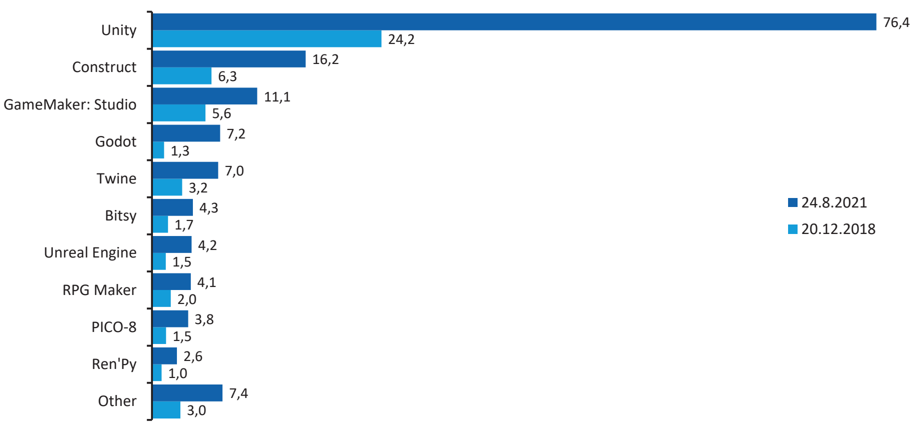
When the engine becomes the driver

Today, modern game engines can comprise some of the most elaborate and complex software ever written. Game engine developers are also targeting an increasingly vast range of hardware platforms and higher-level programming languages, further increasing the decoupling of the different layers of the game industry technology stack. As a consequence, from the perspective of digitalization, the facilitation of the interactions is becoming more and more concentrated in the platform domain.

The game industry is not the first example of a setting where this kind of a decoupling and rebundling of the technology stack has occurred as a result of digital platforms. A decade ago in the mobile phone business, for example, the platformization of the smart phone operating systems transformed the entire industry, leading to a significant change in the bottlenecks and the gate-keeping control points in the value chain. In only a few years, the operating system went from being the most important competitive differentiator between mobile phone manufacturers to becoming almost a commoditized part of the industry's technology stack (Kenney & Pon, 2011; Pon, Seppälä, & Kenney, 2014).

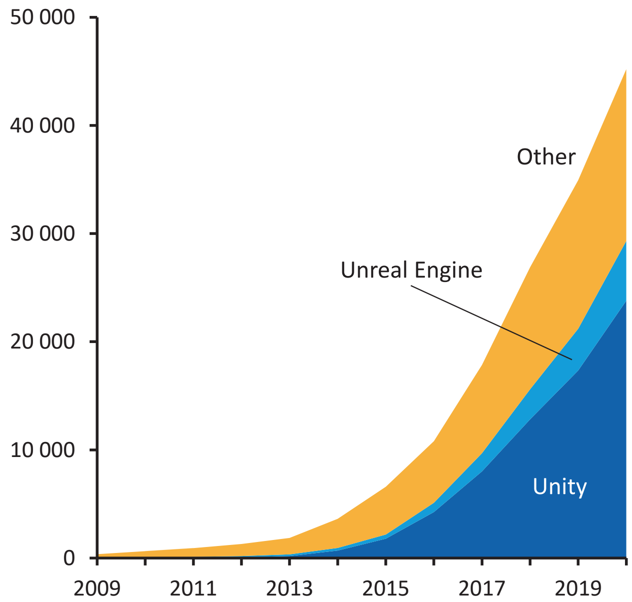
Similarly, just as smart phone devices today are embedded into operating systems instead of *vice versa*, games are now being increasingly embedded into engines instead of engines being embedded into games. In fact, today the majority of game development takes place on top of the few most popular game engine platforms. According to Unity, for example, more than 50 percent of all games across mobile, PC and console domains now utilize the company's game engine, and over 70 percent of the top 1000 mobile games are made on top of the Unity game engine (see Figures 1 and 2).[1] On the game distribution platform Steam, which accounts for 75 % of the global market share of PC games, the two most popular game engines alone account for 39 % of the top 250 most popular games (see Figure 3). The figure is by no means insignificant, considering that one half of Steam's global revenue can be attributed to the top 100 games alone (Zuckerman, 2020).

Figure 1 Number of projects listed in itch.io by game engine (thousands)



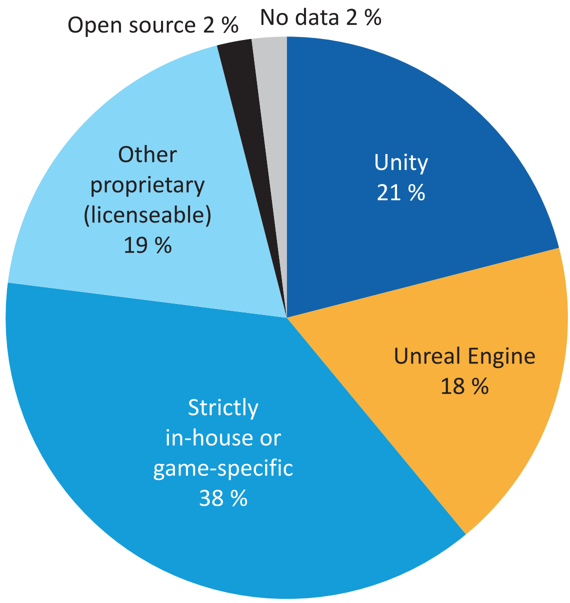
Sources: itch.io.

Figure 2 Total games released on Steam by game engine (cumulative)



Sources: Statista, Steam API data (authors' estimate based on Steam depot files).

Figure 3 Top 250 most popular Steam games by game engine (%)



Sources: Steam API data (authors' estimate based on Steam depot files).

An industry aiming for serious growth

Game engines are a vivid example of how the platforms in the game industry can offer new versatile ways of virtual interaction with data, in a significantly more real-world-like manner than before. The platformization of game engines and their enhanced capability to facilitate virtual interactions have led to the rapid broadening of the scope of their applications in recent years. Instead of pure entertainment, game engines are now increasingly being licensed for non-entertainment purposes such as visualization, training, and scientific exploration in industrial, medical, and military contexts in so-called *serious games*.

To name a few examples, in the construction industry, game engines are now being used to design and explore virtual building information models (BIM) in an interactive real-time manner. As one example, in the aftermath of the fire of the Notre Dame cathedral in 2019, Ubisoft's building information models and game engine were utilized to reconstruct the cathedral's lost historic features. In medicine, game engines are used to train surgeons and to visualize molecular data in the development of new medicine (Gardner, 2018). Respectively, in telecommunications, game engines are employed to simulate 5G wave propagation in real-time (Caulfield, 2021). Also, in the film industry, game engines are enabling new techniques of virtual production where special effect can already be seen in real-time during principal photography (Lappalainen, 2021). For example, Disney's new Star Wars series 'Mandalorian' was filmed and rendered by using Unreal Engine in this very manner (Ball, 2020). The list of applications goes on and on.

Recently, the game industry has increasingly steered its focus towards creating something which, if realized, would enable even more versatile digital interactions. While still existing mostly at the level of a vision rather than reality, these so called *metaverses* are generally described as virtual frameworks that are collectively shared, persistent, synchronous, and interoperable. Comprising more than a mere virtual reality, a metaverse should be understood as a much broader concept, something akin to *system of systems*, a comprehensive infrastructure not tied to any one application or any single individual provider (Ball, 2020).

In academia, systems of systems are typically characterized by five key properties, in the so called 'Maier's criteria'. Firstly, the individual systems must be operationally independent, so that if the system of systems is disassembled, the individual systems can still independently perform in a useful manner. Secondly, the individual systems must be managerially independent, meaning that they are mostly acquired and integrated independently. Thirdly, the individual systems are geographically widely distributed, and can typically readily exchange information but not physical things, such as mass or energy. Fourthly, the system of systems must be capable of emergent behaviour, so that as a collective it can perform higher functions which do not reside in any of the individual systems. And lastly, the system of systems exhibits constant evolutionary development where structure, function, and purpose are continuously added, modified, and removed over time (Boardman, Dimario, Sauser, & Verma, 2006; Maier, 1998; Sage & Cuppan, 2001).

One idea behind the concept of a metaverse is that by building it to be physically based, *i.e.* accurately simulating the laws of physics, material properties, and other such aspects of our physical reality, the metaverse framework can be used for virtually an unlimited scope of purposes in a much more interactive way than in earlier applications. By enabling persistent virtual data objects, digital entities can traverse between digital domains and migrate from one application and industry to another. Furthermore, the idea is that through augmented reality, metaverse objects can also be layered on top of our physical world where they can be interacted with just like any natural object (Caulfield, 2021).

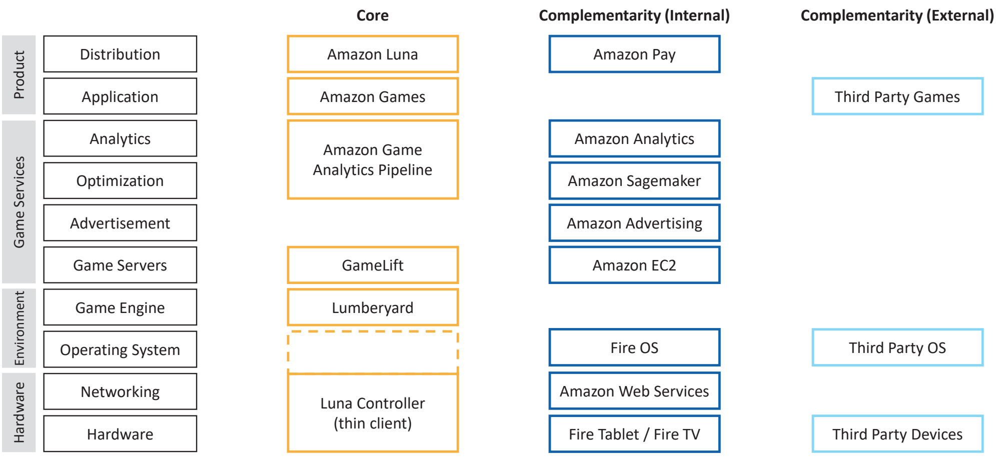
Who weaves the fabric of the new reality?

Should the efforts towards building a system-of-systems-level metaverse platform be successful, the question that naturally follows is who will be the market-makers, the owners of the fabric of this new reality, and how will this affect the value capturing ability across industries and geographies?

Unity, the provider of the most popular game engine at the moment, has recently given indications of its plans to engage in metaverse development (Gabriele, 2020; VentureBeat, 2021; Parisi, 2021). Similarly, Epic Games has also expressed a desire to develop Fortnite, one of the most popular games based on its engine, into a metaverse platform (Ball, 2020). Many others affiliated with the game industry have also expressed similar goals. For example, the graphics hardware manufacturer NVIDIA recently launched a new engine platform named Omniverse, which the company says is “aiming for universal interoperability” (nVidia, 2021). Facebook—more recently known as Meta—has voiced its desire to transform itself from a social network company into a metaverse company, with the help of its Oculus VR and Horizon virtual meeting space technologies (Newton, 2021). Amazon, Microsoft, Epic Games, and Valve have also all been increasing their capabilities and resources along a similar tangent (Ball, 2020).

Additionally, many of the incumbent digital platform giants of today, such as Google, Apple, and Tencent, already possess strong monolithics which they could leverage for market dominance in the wake of the transformation of the game industry. For example, Google’s ecosystem already facilitates strong capabilities in streaming, live ops, and artificial intelligence—all of it supported by immense amounts of data. Similarly, Amazon has been systematically increasing its capabilities on all levels of the game industry’s technology stack, including the Luna Controller thin client gaming hardware, the Lumberyard game engine, and Amazon Luna—a Netflix-style distribution platform for streaming games over the internet (See Figure 4).

Figure 4 The game industry technology stack of Amazon



Sources: Authors.

By having a stake of ownership in the most prominent game engine providers, the Chinese platform giants have also been able to tap into the European and American markets through this horizontally expanding platform layer in the game industry. For example, Tencent has formerly acquired 40 % of Epic Games, the provider of Unreal Engines (Kain, 2021).

While Western digital multi-sided platform giants have had the tendency to grow and evolve more vertically across the technology stack, the Chinese platform giants have more experience in building and expanding their platform businesses horizontally across industries, through their so-called *platform business group strategy* (see *e.g.* Jia, Kenney, Mattila, & Seppälä, 2018). Whether the Western platforms will be able to adapt to this horizontal model efficiently remains an open question, and a factor in how the platform landscape may become reconfigured in the wake of the metaverse development.

How should Finland prepare?

Finland has a long history of successful video game development. In recent years, some Finnish game studios have arisen amongst some of the most important companies for the Finnish economy in terms of GDP contribution (Ali-Yrkkö, Seppälä, & Mattila, 2016). In recent international comparisons, Finland has placed amongst the top three game developer countries in Europe by turnover, making it one of the most attractive game industry hubs in the world today (Neogames, 2019). In this regard, the increasing trend of platformization in the game industry and the metaverse development raise several considerations for Finnish business strategy and public innovation policy.

The vision of an interoperable system of systems is not entirely a new one to the industrial sector, of course. A similar idea has been baked into many earlier industrial concepts, such as 'internet of things', 'industrial internet', 'digital twin', and so on (Porter & Heppelmann, 2014). Many problems have been identified in research that stand in the way of this development, as earlier efforts have struggled to establish wide-scale integration. (Tähtinen, 2018). In the light of these prior difficulties, one might ask, what separates the game industry platforms from earlier efforts towards system-of-systems-level interoperability.

Despite the momentum, the concept of the so-called 'data economy' has so far remained ambiguous in regard to its significance to most companies and industries, and the Finnish economy in large (Nikander, Mattila, & Seppälä, 2018; Tähtinen, 2018). One of the key considerations is whether the new wave of game industry platforms and the metaverse development can crystalize the concept the so-called 'data economy' and its significance to industries through their enhanced interaction and virtual economies. By making the benefits more understandable through the increased capacity for real-world-like interaction, and by providing a complete workspace with a wide range of development tools and interfaces, the development could take a different trajectory from earlier attempts.

Furthermore, one of the key problems with enabling interoperability of industrial data has been the absence of a platform which could conveniently facilitate the incentive structure for providing data monetizing its use. The game industry's expertise in building virtual economies puts it in a unique position to establish data product markets, potentially unlocking industrial data interoperability. For example, could the platformization of serious games provide a way for industrial companies to tap into game industry's virtual economies? Can game industry platforms make contractual arrangements regarding data ownership and data governance easier than before? Could such a development provide an incentive for industrial players to defuse their horizontal barriers of data product interoperability?

From the perspective of innovation policy, it is important to understand the general applicability of this new enhanced capability of facilitating interactions which game industry platforms and the metaverse development have to offer. As a consequence, companies and policy makers alike should seek to increase their understanding on which industries will be affected by this development in the near future, in what capacity, and under what kind of a timeframe.

Thirdly, if the game industry platforms and the metaverse development are successful in system-of-systems-level integration across industries, one key consideration is whether the game industry companies can challenge the current digital platform incumbents, such as Google, Amazon, Facebook and Apple, as the providers of the next generation of digital infrastructure. As discussed above, the digital platform incumbents have also been actively bolstering their capabilities in the game industry's technology stack through mergers and acquisitions. From the standpoint of competition policy and antitrust, it is important to pay special attention to these types of acquisitions and consider how these capabilities are being fused into the service offerings of the current incumbent platform giants.

As the game content creation has become increasingly separated from the development of the game engines, smaller game studios have become more and more dependent on the game industry's digital platforms. In order to protect the Finnish game industry and the Finnish national economy from falling victim to predatory innovation in this domain, careful consideration should be exercised on how to keep the next generation of digital infrastructure from slipping through the fingers of the Finnish innovation ecosystem. A key consideration for strategy and policy in this regard is how the resources, the knowledge and the tools already present in the different settings of this problem domain be leveraged against one another. Furthermore, companies and policy makers should seek to understand what kinds of resources, protocols, and regulative frameworks will be required to foster new businesses and industrial growth in these new digital infrastructures in the near future.

Endnote

[1] https://unity.com/our-company

References

Ali-Yrkkö, J., Seppälä, T., & Mattila, J. (2016). *The Role of the Largest Companies and Their Value Chains in the Economy*. Retrieved from http://pub.etla.fi/ETLA-Raportit-Reports-53.pdf

Ball, M. (2020). The Metaverse: What It Is, Where to Find it, Who Will Build It, and Fortnite. Blog post. https://www.matthewball.vc/all/themetaverse (Accessed 4.11.2021).

Boardman, J., Dimario, M., Sauser, B., & Verma, D. (2006). System of Systems Characteristics and Interoperability in Joint Command and Control. Ft. Belvoir, VA: 2nd Annual System of Systems Engineering Conference.

Caulfield, B. (2021). What Is the Metaverse? *NVIDIA Blog*. Retrieved from https://blogs.nvidia.com/blog/2021/08/10/what-is-the-metaverse/ (Accessed 4.11.2021).

Gabriele, M. (2020). Unity is Manifesting the Metaverse. The Generalist. https://mariogabriele.medium.com/unity-is-manifesting-the-metaverse-b96d1639c95c (Accessed 4.11.2021).

Gardner, E. (2018). Unreal Engine: From Gaming to Ground-Breaking Cures. Pharma Technology Focus 77(12). https://pharma.nridigital.com/pharma_dec18/unreal_engine_from_gaming_to_ground_breaking_cures (Accessed 4.11.2021).

Jia, K., Kenney, M., Mattila, J., & Seppälä, T. (2018). *The Application of Artificial Intelligence at Chinese Digital Platform Giants: Baidu, Alibaba and Tencent*.

Kain, E. (2021). 'Fortnite' Is Calling It Quits In China. Forbes. https://www.forbes.com/sites/erikkain/2021/11/02/fortnite-is-calling-it-quits-in-china/ (Accessed 4.11.2021).

Kenney, M., & Pon, B. (2011). *Structuring the Smartphone Industry: Is the Mobile Internet OS Platform the Key?* (SSRN). SSRN. https://doi.org/10.2139/ssrn.1851686

Lappalainen, E. (2021). Se toinen urahaave. HS Visio 20.8.2021. https://www.hs.fi/visio/art-2000008188748.html (Accessed 4.11.2021).

Maier, M. W. (1998). Architecting Principles for Systems-of-Systems. *Systems Engineering, 1*(4), 267–284. https://doi.org/10.1002/(SICI)1520-6858(1998)1:4<267::AID-SYS3>3.0.CO;2-D

Neogames Finland ry. (2019). The Game Industry of Finland Report 2018.

Newton, C. (2021). Mark in the Metaverse. Facebook's CEO on why the social network is becoming 'a metaverse company'. The Verge podcasts. https://www.theverge.com/22588022/mark-zuckerberg-facebook-ceo-metaverse-interview (Accessed 4.11.2021).

Nikander, P., Mattila, J., & Seppälä, T. (2018). Kehittymättömät datamarkkinat johtavat Internetin keskittymiseen. In *Digibarometri 2018* (pp. 31–37). Helsinki.

nVidia (2021). Omniverse Platform Overview. https://docs.omniverse.nvidia.com/plat_omniverse/plat_omniverse/overview.html (Accessed 4.11.2021).

Parisi, T. (2021). The Seven Rules of the Metaverse. Blog post. https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c (Accessed 11.11.2021).

Pon, B., Seppälä, T., & Kenney, M. (2014). Android and the demise of operating system-based power: Firm strategy and platform control in the post-PC world. *Telecommunications Policy*, 1–13. https://doi.org/10.1016/j.telpol.2014.05.001

Porter, M., & Heppelmann, J. (2014). How Smart, Connected Products are Transforming Competition. *Harvard Business Review*.

Sage, A., & Cuppan, C. (2001). On the Systems Engineering and Management of Systems of Systems and Federations of Systems. *Information-Knowledge-Systems Management*, *2*(4), 325–345.

Still, K., Seppänen, M., Seppälä, T., Suominen, A., Valkokari, K., & Korhonen, H. (2017). Alustatalous on vuorovaikutustaloutta. Helsinki: ETLA Economic Research.

Tähtinen, T. (2018). *How to Turn IIoT Software Solutions into Platforms: A Case Study of Three Existing Solutions*. Aalto University. Retrieved from http://urn.fi/URN:NBN:fi:aalto-201809034820

VentureBeat (2021). How Unity is positioning itself to democratize the metaverse. GamesBeat blog. https://venturebeat.com/2021/01/27/how-unity-is-positioning-itself-to-democratize-the-metaverse/ (Accessed 4.11.2021).

Zuckerman, A. (2020). 75 Steam Statistics: 2020/2021 Facts, Market Share & Data Analysis. CompareCamp. https://comparecamp.com/steam-statistics/ (Accessed 4.11.2021)

ARTICLE 7

Platform-Dependent Entrepreneurs:

Power Asymmetries, Risks, and Strategies in the Platform Economy

Donato Cutolo[a]
Martin Kenney[b]

Affiliations:

[a] University of Bologna, Italy
[b] University of California, Davis, USA

Original publication:

Cutolo Donate and Kenney Martin, Academy of Management Perspectives, 2021, Vol. 35, No. 4, 584–605.

Abstract

Online digital platforms organize and mediate an ever-increasing share of economic and societal activities. Moreover, the opportunities that platform-mediated markets offer not only attract enormous numbers of entrepreneurs, but also support the growth of entire ecosystems of producers, sellers, and specialized service providers. The increased economic and business significance of digital platforms has attracted an outpouring of studies exploring their power dynamics and general impact. This research has largely overlooked the power imbalance that entrepreneurs experience as members of the platform ecosystem and provided little guidance on how these far more numerous firms should compete. Drawing upon Emerson's power-dependence theory, we show that the power asymmetry at the heart of the relationship between the platform and its ecosystem members is intrinsic to the economics and the technological architecture of digital platforms. We undertake a conceptual analysis of the sources of this power, and we unravel the novel component of risks that emanate from this imbalance. Our analysis suggests that the conditions of engagement for platform entrepreneurs are so different from traditional entrepreneurship that these entrepreneurs are more usefully termed "platform-dependent entrepreneurs" (PDEs). Further, we explore the strategies that PDEs are developing to mitigate their dependence. Finally, our study provides a framework for policy makers that are considering regulating platform-organized markets.

Keywords

Entrepreneurship, Platforms, Digital markets, Platform dependent entrepreneurship, Strategy, Ecosystems, Complementors

1 Introduction

The role online digital platforms play in controlling commerce and communication means that entrepreneurs and, indeed, a substantial portion of all businesses have to navigate a world where platform structure the reality (Cennamo, 2019; Cusumano, Gawer, & Yoffie, 2019). Because of network effects and winner-take-most aspect of these markets (Gawer and Cusumano 2002, Cennamo & Santalo, 2013), successful digital platforms have coalesced into powerful economic intermediaries. As a result, the economy is being (re)structured by platform firms and participation in these ecosystems has become vital for many businesses' existence and growth (Kenney & Zysman, 2016; Parker, Van Alstyne, & Choudary, 2016). To illustrate, it is possible to ask whether an organization that cannot be found through Google Search exists, a restaurant can afford to ignore Yelp, an online business can ignore Amazon – that now is estimated to control approximately 40% of all online sales, or hotels can afford not to rent through the online travel agencies. In October 2019, of the ten most valuable firms in the world, seven were digital platform firms, quite simply because the stock market believes they are in a position to capture an enormous share of the world market's total value.

Entire constellations of producers, sellers, and specialized service providers have emerged to earn their livelihoods through these platforms. Digital platforms such as Amazon, eBay, Etsy, Facebook, Google, Instagram, Yelp, and YouTube, among others, make it easier than ever to build a business and generate income, offering entrepreneurs access to large-scale markets and a variety of incentives to populate their platform ecosystems (Ghazawneh & Henfridsson 2013; Yoffie & Kwak, 2006). We will demonstrate that the conditions of selling or providing services through a platform are so different from traditional entrepreneurship that entrepreneurship actualized through an online digital platform can usefully be termed "platform-dependent entrepreneurship"[1].

This paper builds upon and extends the recent outpouring of research on platform entrepreneurship (Nambisan, 2017; Kapoor & Agarwal, 2017; McIntyre & Srinivasan, 2017; Eckhardt, Ciuchta, & Carpenter, 2018). When exploring the nature of entrepreneurship conducted on platforms, extant studies have emphasized the peculiarity of this context, which is characterized by network effects and winner-take-most outcomes that pose novel challenges for platform-dependent entrepreneurs (PDEs) selling through these platforms (Gawer & Cusumano 2002). While recognizing the tremendous new business opportunities created by online platforms, surprisingly, little attention is given to the power relationship between PDEs and platform owners. As members of a platform ecosystem, PDEs experience a great power imbalance in relation to the platform owners, who can unilaterally enforce changes in the competitive conditions on the platform (Kapoor & Agarwal, 2017; Wen & Zhu, 2019). Although recognition of this power imbalance is growing (Miric, Boudreau &

Jeppesen, 2019; Nambisan & Baron, 2019), there has not been a comprehensive exploration of the power dynamics faced by PDEs.

The purpose of this paper is to advance a perspective that extends and enriches our understanding of how the power asymmetries inherent in digital platforms influence and restructure entrepreneurship. We first define and describe the nature of platform entrepreneurship, showing how extant research fails to provide an adequate account of the relationship between platform owners and PDEs. Next, we rely on power-dependence theory (Emerson, 1962) to illustrate how the power imbalance in this relationship arises from the technological and economic dynamics of digital platforms and is intrinsic to platform architecture and design. By detailing the sources of power, we show that the entrepreneurial process, which is already characterized by high risk, is made even more precarious by being dependent upon a platform. In this regard, we show that unique and pervasive risks stem from this dependence. Of course, PDEs have introduced strategies that, while limited in efficacy, can provide some countervailing power. In the discussion, we reflect on how and why this changes the theory and practice of entrepreneurship, emphasizing policy implications, and promising areas for future research.

2 Theoretical background: Power asymmetry in platform-dependent entrepreneurship

Platforms have been defined in a variety of ways (Baldwin & Woodward, 2009; Parker et al. 2016; Evans, Hagiu, & Schmalensee, 2006). We adopt Gawer's (2014: 1240) definition "that platforms are evolving organizations or meta-organizations that: (1) federate and coordinate constitutive agents who can innovate and compete; (2) create value by generating and harnessing economies of scope in supply or/and in demand side of the markets; and (3) entail a modular technological architecture composed of a core and a periphery." Any platform thus implies the presence of a group of actors, or complementors, that supply complementary products and services that generate value for the core platform business (Gawer & Cusumano, 2002; Parker et al., 2016). Complementors join a platform's ecosystem for a variety of reasons (Boudreau & Jeppesen, 2015). More recently, considerable academic interest has focused on the complementors that join a platform ecosystem with entrepreneurial intent (Nambisan, 2017; Eckhardt, et al., 2018; Nambisan & Baron, 2019). While there are non-profit platforms, the phenomenon that we address are those where both the platform owner and the complementors are entrepreneurs producing goods or services for income or for-profit entities intent upon maximizing their income.

To explain entrepreneurial action on a platform, academic attention has focused upon the impact of digital technologies on entrepreneurship-related concepts, addressing how the technological dimension of digital platforms operates to define entrepreneurial opportunities, processes and outcomes. Nambisan (2017) highlights the need for developing theory that addresses the relationship between digital technologies and entrepreneurship, as well as how digital platforms alter the uncertainty inherent in the entrepreneurial process—since conducting entrepreneurship on a digital platform implies more blurred boundaries and dispersed agency. Building on that, von Briel, Davidsson, and Recker (2018) emphasize the central role digital platforms have as enablers of entrepreneurial opportunities, dissecting platforms' influence on the agency and boundaries of venture creation at different stages of the entrepreneurial process. Exploring entrepreneurship in digital platform-organized markets must consider its unique features, such as generativity (Zittrain, 2008), technology affordances (Autio, Nambisan, Thomas & Wright, 2018), and openness (see Nambisan, Siegel and Kenney (2018).

In addition to the central role of digital technologies, scholars suggest focusing on digital platforms as a novel and unique setting for entrepreneurship. Digital platforms orchestrate entire ecosystems of value creation and exchange, opening new spaces and channels where entrepreneurs can create new firms and operate (Nambisan, 2017; Jacobides, Cennamo, & Gawer, 2018; Cusumano et al., 2019). A deep understanding of the entrepreneurial context serves multiple purposes from a theoretical standpoint, since the character of entrepreneurship, as well as the actions and outcomes of any entrepreneurial effort, depend on the rules, threats, and opportunities framing its context (Autio et al., 2014). Thus, understanding platform-dependent entrepreneurship requires explicating the context for this entrepreneurship and explaining the reasons for the dependence that emerges.

Platform-based entrepreneurship differs substantially from traditional entrepreneurship. As Nambisan and Baron (2019) point out, PDEs simultaneously fill two roles. First, PDEs operate businesses pursuing goals, with the platform as an intermediary. However, to the platform owner, the PDEs are complementors, whose existence is only important if it adds value to the platform. Consequently, entrepreneurial processes and outcomes are conditioned by the dynamics determined by membership in a digital platform ecosystem. To illustrate, Eckhardt et al. (2018) find that in an app store, digital platforms provide ecosystem members with information regarding the commercial feasibility of their products, thereby influencing their propensity to commercialize their software programs. In contrast, McIntyre and Srinivasan (2017) adopt a network perspective to illustrate how entrepreneurial success on a digital platform is intertwined with the fast-paced competitive dynamics that characterize digital platforms and their ecosystems.

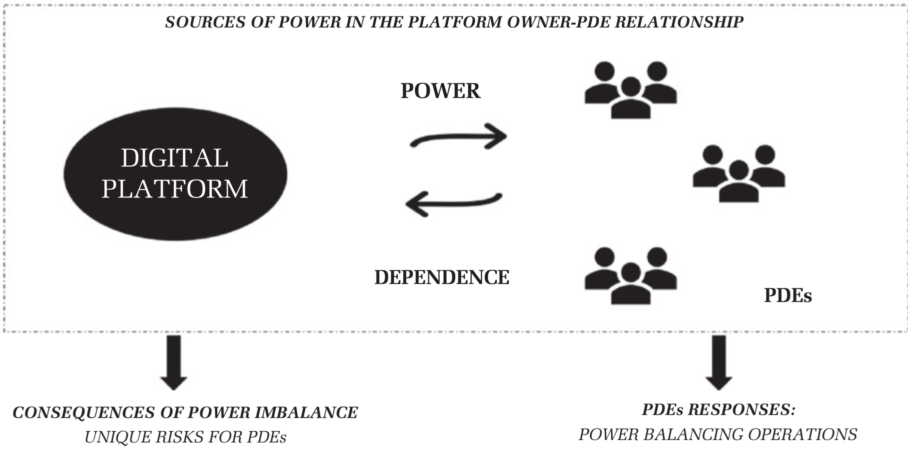
There have been significant efforts to integrate different literatures to articulate the theoretical foundations for platform entrepreneurship (Nambisan et al., 2019), while also recognizing the uniqueness of digital platforms as entrepreneurial contexts (Kapoor & Agarwal, 2017; Eckhardt et al., 2018; Nambisan & Baron, 2019). And yet, there has been less analysis of the relationship between entrepreneurs as ecosystem members and platform owners. In a recent paper, Jacobides et al. (2018: 2258) observed that the relationships between platform owners and ecosystem members differs from other inter-firm relationships as they "do not fit into the classical firm-supplier relationship, Porter's (1980) value system, or a firm's strategic networks; neither are they integrated hierarchies." Previous studies nearly always unquestioningly postulate that the relationship between the platform owner and complementor is based upon the shared objective of providing value to customers (Nambisan & Baron 2013; Adner, 2017; Jacobides et al., 2018), and thus, accept that the actors "depend on each other and share a common fate" (Tiwana, Konsynski, & Bush. 2010: 52). Others go further, depicting the relationship as a partnership (Gawer & Cusumano 2014; Zhu & Liu, 2018; Wang & Miller, 2019). Remarkably, these authors do not reflect upon what "partnership" means given the fundamental asymmetry between the platform owner and the PDE (Boudreau & Hagiu, 2009).

Platform ecosystems are not fully hierarchically controlled and their participants are separate organizations (Jacobides et al., 2018). Despite this lack of direct control through ownership, platform owners can "impose rules and constraints, create inducements and otherwise shape behaviors" (Boudreau & Hagiu 2009: 3). Nambisan and Baron (2013: 1073) correctly observe that the complementors "surrender part of their autonomy and independence" and align their businesses with the goals of the platform owner. As the intermediary between potential customers and providers, there is an asymmetric power relationship that affects the entrepreneurs' return, though the conditions of that relationship may change. While there is ample evidence of power asymmetry (Kapoor & Agarwal, 2017; Miric et al., 2019; Nambisan and Baron, 2019), there has been little consideration of the mechanisms through which platform owners wield power and of the consequences of this power imbalance. This gap in the literature leaves unaddressed critical questions regarding changes to our understanding of entrepreneurship, when increasing portions of the economy are organized by platforms.

3 Sources of power in the platform owner-PDE relationship

Platform power has already been an object of study, but scholars have mainly focused on the market power of platform firms and its consequences for competition (Eisenmann, Parker, & Van Alstyne, 2011; Khan, 2016). Despite the clear power asymmetries within the platform owner-entrepreneur relationship, their effect is explored only in passing (McIntyre & Srnivsan, 2017; Gerwe & Silva, 2018; Jacobides et al, 2018). In order to understand how this power asymmetry influences and transforms entrepreneurship, it is necessary to elucidate the sources of this power.

Figure 1 Platform Dependent Entrepreneurship: sources of platform power, consequences of power imbalance, and PDEs' mitigation responses



Adapted from Emerson (1962).

Power is an attribute of the relationship between actors rather than a consequence of their individual characteristics (Cook et al., 1983; Emerson, 1962). The central proposition of power-dependence theory (Emerson, 1962) is that, within any relationship, power stems from 1) control over valued or needed resources by others and 2) the availability of alternative sources for these resources. Figure 1 illustrates how Emerson's power-dependence lens frame our exploration of platform dependent entrepreneurship. We first outline that platforms' power is rooted in some of the techno-economic features of digital platforms, and we show that both the incentives developed by the platform to attract PDEs and the mechanisms designed to ensure their commitment exacerbate power imbalance within the PDEs-platform relationship. We then proceed to explore the consequence of this asymmetric distribution of power detailing the unique risks that PDEs face and we illustrate the balancing responses developed by PDEs to mitigate that power.

3.1 Techno-economic bases of platform power

The power imbalance in the platform owner-PDE relationship originates in the first instance from the digital nature and the peculiar dynamics of digital platforms. The fundamental source of power for online digital platforms is, of course, the value provided to its users. As an intermediary, a platform connects different groups of users (economists term these as “sides” of a platform). Through connecting users and presumably offering them sufficient value to retain them and attract others, the platform can generate positive network effects, attracting yet more users from both the sides (Parker et al. 2016). These multisided platform dynamics in fact exhibit direct and indirect network effects (Parker & Van Alstyne, 2005), which represent the main driving force of platforms’ value and market share (Cennamo & Santalo, 2013). Often, the platform can even attract entirely new groups (sides) of users. For example, YouTube solved what has been termed the “chicken-and-egg” problem by seeding a few videos that attracted viewers and providing tools for viewers to easily upload videos and embed them in other pages, thereby attracting attention to YouTube. As these two sides grew, YouTube began attracting advertisers, which were a new group of users and side of the platform. YouTube or any platform’s success is predicated upon igniting positive feedback loops or what Cusumano et al. (2019) characterize as “rapid nonlinear growth”.

The strength of these network effects is such that it can easily lead to high levels of market concentration, thus the successful platforms are often winner-take-most/all (Parker & Van Alstyne 2005; Eisenmann et al. 2011). Winner take-all/most outcomes are at the heart of platform value creation and capture (Amit & Zott, 2001) and they often result from competition between platforms. The belief or hope that a platform could disrupt existing competitors justifies even a willingness to suffer financial losses to capture the market.

In this process, platforms provide both consumers and producers with incentives and benefits to join and maintain their association with the platform ecosystem, and since these benefits are amplified by the increasing returns associated with network effects, they often result in lock-in (Arthur 1989; Ozalp, Cennamo, & Gawer, 2018). [2] The lock-in of platform users is reinforced by other attributes of platforms, such as the long-tail effect, which refers to the fact that dominant platforms have “everything” including the most obscure items. The ability to find anything means that customers have no need to search elsewhere, increasing the chances that they will engage in repeated transactions, thereby strengthening lock-in. To illustrate, Amazon, through its Marketplace, has “everything.” Estimates vary, but one website suggests that in April 2019 Amazon carried 128 million unique items; of which 44.2 million were books -- of course, there also might be many offerings of any particular item[3]. YouTube has a similarly enormous numbers of videos, as it is estimated that 500+ hours of video are uploaded to YouTube every minute[4]. Almost invariably, all the items in the long tail are provided by PDEs, so they bear the cost of proposing the item.

Once lock-in has been achieved, there are very few alternatives and PDEs will inevitably be in a position of dependence. Moreover, PDEs must optimize their operations for the dominant platform, thereby deepening their lock-in. For example, there is an entire industry providing “search engine optimization” services, which does nothing more than design websites to be found, catalogued, and ranked highly by Google. These basic dynamics of successful digital platforms' strategy have implications for the way the market operates and for the power imbalance experienced by PDEs.

In addition, several features of digital technologies combine to explain how digital platforms develop their power. As intermediaries, platforms provide a digital infrastructure that lowers search and transaction costs for both sides of the market and improves the match between the parties (Baldwin & Woodward, 2009). Because the platform is built from software, this infrastructure can be easily altered and reconfigured. For example, the *Search Engine Journal* (2019) finds that Google makes thousands of minor changes each year, and, less frequently, major changes to its search algorithms – presumably each of these changes is in Google's interest. The ability to control and alter the technical infrastructure upon which users participate and contribute to the ecosystem allows platform owners to influence other participants, directly or indirectly.

Since all actions on platforms are digital, they are all recorded, and thus giving the platform a panoptic view of the activities of all users (Zuboff 2019). Given its central position, the platform can decide what information to provide to which users and, of course, what will not be provided. This ability to analyze, recombine, and manipulate data and information allows the platform to influence attention and actions (Gerwe & Silva, 2018). As an example, collection of the online actions of each user allows a platform to serve "individualized" content to each user.

This capability to control data, direct attention and orient behaviors is fundamental to a platform's relationship with PDEs and it places platform owners in a position of considerable power as they can use it in their own favor, for instance promoting their own offerings. To illustrate, according to a recent analysis conducted by the *New York Times*, in more than 700 searches in Apple's online store, Apple ranked its apps first over competitors. For example, some searches for term "podcast" returned 14 Apple apps before showing results from other companies[5].

3.2 Platform incentives and resources to PDEs

Particularly in the initial stage of a platform's life, when network effects are minimal, it is often necessary to provide significant and money-losing incentives to attract PDEs and/or consumers in an effort to "tip the market" (Arthur 1989; Shapiro & Varian 1998). Attracting entrepreneurs is critical and offering attractive terms is vital, as the platform is often competing against others (Gawer & Cusumano, 2002). However, the incentives and the resources provided by digital platforms to attract and cultivate their relationship with PDEs have a contradictory effect. We illustrate how the very same benefits associated with joining a platform becomes the sources that sustain and reinforce the power asymmetry in the PDEs-platform relationship.

3.2.1 Access to customers

For those selling goods or services, the fundamental benefit of using a platform—whether it be advertisers paying for search advertising, Etsy sellers, or IOS developers—is customer access. Digital platforms are "matchmakers" (Evans & Schmalensee, 2016:1) and this refers to a platform's ability to match buyers and sellers or service providers, and to reduce discovery and transaction costs (Evans, Boudreau, & Hagiu, 2009; Baldwin & Woodward, 2009). A platform's market ranges from global (online sales, e.g., Amazon marketplace) to extremely local (e.g., Yelp! for locating a restaurant), but in aggregate, their scale is enormous (Cennamo & Santalo, 2013). To cope with the size of these markets, platforms offer classification systems, e.g., tags, product categories and more, that make discovery of far-flung sellers possible, thereby reducing discovery costs and creating new spaces for entrepreneurs. Control over access to customers is the fundamental first dimension of a platform's power, as it directly affects the most valued and essential resource for the PDEs: access to the market. All things being equal, as a direct consequence of network effects and winner-take-most dynamics described in section 3.1, the greater the concentration of users/customers on a single platform in a particular market, the greater the power over its PDEs.

3.2.2 Provision of boundary resources

The fundamental problem faced by every platform is to attract different groups of actors, namely, at a minimum, providers of a desired good or service (PDEs) and users/customers (Gawer & Cusumano, 2014). To attract these actors, a platform provides them with tools, (such as, software development kits, application-programming interfaces (APIs), marketing and sales information, training, templates, manuals, technical support and other resources (Boudreau & Hagiu 2009; Ghazawneh & Henfridsson 2013; Yoffie & Kwak, 2016; Eckhardt et al., 2018). These are provided to all sellers and facilitate use of the platform (Eisenhardt & Martin, 2000). These resources are the affordances that reduce both the entry barriers and scaling costs for PDEs (Eckhardt et al., 2018; Nambisan et al., 2018).

The provision and control over these resources grant platform owners considerable power by virtue of two mechanisms. First, the boundary resources generate power by forcing complementors to make asset-specific investments as a condition for participation (Eckhardt et al., 2018). The greater the investment is—which is often cumulative due to platforms' ranking and reputation systems—the greater the power platform owners accrue (Luca & Zervas 2016). In other words, platforms attempt to create lock-in and limit the possibility for PDEs to pursue economic interests outside the platform. Second, boundary resources are designed to control actions on the platform (Ghazawneh & Henfridsson 2013), as they specify the parameters of permissible action (Eaton et al., 2015: 220).

3.2.3 Platform governance

The platform owner is responsible for the functioning of the ecosystem through the provision of modular architecture and by setting the rules of engagement for actors (Gawer & Cusumano, 2002, Wareham, Fox, & Giner, 2014). As ecosystem curators, platform owners must coordinate their PDEs to prevent dysfunction (Thies, Wessel, & Benlian, 2018; Jacobides et al. 2018): platform governance encompasses decision rights partitioning, control mechanisms, and pricing policies (Tiwana et al., 2010). In other words, platform owners act as private regulators who are expected to reduce negative externalities created by ecosystem members in order to maximize the value for the system as a whole (Boudreau & Hagiu, 2009; Evans, 2012). The profit of the platform owner and the value of the ecosystem are directly linked, and insufficient control over opportunistic behaviors by PDEs can degrade the ecosystem and even result in a platform's failure (Täuscher & Kietzmann, 2017). Platforms are thus strongly incentivized to perform a regulatory role, and they have a large set of control mechanisms to do so, (Evans, Hagiu, & Schmalensee, 2006) including data-driven technologies, such as algorithmic recommendation and reputation metrics, gatekeeping, and exclusion from the ecosystem (Curchod, et al., 2019).

Platform governance also sustains the power asymmetry between owner and PDEs. The ultimate source of a platform's power is its ownership of a digital "space" and within this digital space, the owner has the right to set and change any parameter—barring violation of the law. This power is expressed in two ways:

First, there are the "hard" technical components that are the core of the platform. These include the data, algorithms and boundary resources provided, including software development kits (SDKs) and APIs. These frame actions, e.g., only a video with such-and-such specifications can be uploaded on YouTube, etc. (Ghazawneh & Henfridsson 2013; Eaton et al., 2015), or only particular types of data can be inputted to or extracted from the platform. To illustrate, before 2016 Uber did not include a timer (hard-coded in the driver's app) that counted down the five minutes a driver had to wait before being able to leave and collect the cancellation fee. Prior to including the timer, the drivers had to estimate the time of their wait, because if they left and the tardy passenger complained, the driver might lose the cancellation fee particularly because Uber "recommended" that drivers wait ten minutes. The timer was only implemented for all Uber services after the Federal Trade Commission opened hearings on the matter (Rosenblatt, 2018: 120-121). Implementing the wait timer created transparency, which provided drivers with protection. In another case, as Rosenblatt (2018: 122-123) shows, in 2016 Uber implemented "up-front pricing", which allowed a rider to know the price in advance. However, prior to up-front pricing a passenger could wait in the car and compare what they paid with what the driver received. With the new system, Uber instituted a delay so that the driver and rider could no longer make this comparison. These are simply anecdotal illustrations of the more general point, which is that the goals of a platform can be hard-wired into its technical components—and in each case, these software implementations were undertaken without discussion with the affected parties

Second, to operate effectively, many "soft" components, such as rules, principles of community, etc. are designed to channel and control the actions of the actors. These provide guidance on acceptable behavior that include types of content, legitimate action on the platform, etc. These soft components can be powerful because they are vague and thus provide broad parameters for platform action. The principles of community have often been reinterpreted to prohibit previously approved actions, such as, when YouTube demonetizes videos posted prior to the reinterpretation of its principles[6].

Quite simply, platforms can unilaterally set the terms of engagement for PDEs, and this power is intrinsic to platform design, technological architecture, and terms and conditions of use. Starkly put, platform users have two choices—accept the technical and contractual conditions or cease using the platform.[7]

4 Consequence of the power imbalance: Unique risks for PDEs

The power asymmetry at the crux of the owner-PDE relationship can be understood as an asymmetric distribution of dependence between the actors (Emerson, 1962). Although the platform-PDE relationship has some resemblance to other asymmetric inter-organizational relationships characterized by a strong power imbalance, such as those documented in the literature on global value chains (Gereffi, Humphrey, & Sturgeon, 2005; Katila, Rosenberger, & Eisenhardt, 2008; Yamin et al., 2015), the PDE relationship has the following features that make it fundamentally different.

First, despite the fact that imbalances in other inter-firm relationships are also predicated upon resources uniquely provided by a more powerful partner (Katila et al., 2008), traditional supplier relationships are better balanced. First, the supplier often has multiple channels from which they can select, prosaically, Walmart, Costco, and Target (Yamin et al., 2015). Due to winner-take-most dynamics in platform-based markets (Cennamo & Santalo, 2013), the platforms are often quasi-monopolies, leaving few alternatives for PDEs. To illustrate, Apple and Google account for 97% of the mobile operating system market share ex-China. Thus, alternatives are virtually non-existent even for large firms.

This extreme concentration results in the contractual obligations regulating the platform-PDE relationship differing markedly from the traditional supplier-buyer relationship. For nearly all platform users, the terms and conditions of participation are non-negotiable. Even powerful actors, such as Spotify, have found it nearly impossible to demand better terms from the Apple app store. Whereas a traditional supplier usually signs a long-term agreement that normally includes protections for both sides, the contracts signed with platforms invariably permit unilateral changes and with little or no notice. Such changes may alter the terms and conditions, various algorithms, website structure, and profit margins. PDEs can petition the platform to rescind or alter its decisions, and as Eaton et al. (2015) shown, the platform, may, at its own discretion, relent. Moreover, unless the contract violates the law, there is rarely any legal recourse. Finally, by its very nature, in contrast to supplier relationships, the transactions over the platform are not transactions with the platform.

Conceptualizing platform-dependent entrepreneurship as a unique power-dependence system allows a better understanding of the actions and outcomes for the actors involved. It is axiomatic that entrepreneurs face not only everyday business risk but also uncertainty (Schumpeter, 1942; Knight, 1921). However, entrepreneurs building a business on a platform face unique risks that emanate directly from the inherent nature of platforms and the power they wield over members of their ecosystem. For instance, as an intermediary, a platform separates PDEs from customers. Platform owners have the ability to enter into the market space of their PDEs and, immediately, benefit from deep visibility into their now competitor's business. The power-dependence asymmetry provides the platform with the ability to shift the competitive conditions in its favor and to overcome resistance to its actions, and appropriate more value from the member of its ecosystem. Due to the winner take most dynamics, alternatives decline or disappear. As a result, the terms of engagement shift decidedly in favor of the dominant platform.

4.1 Separation from Customers

The relationship between a seller and their customer is fundamental and vital for discovering customers' needs and benefiting from user-led innovation (von Hippel, 1988). Because the platform is the intermediary between the actors transacting on the platform, it is in the platform's interest to keep the sides estranged. As such, the platform channels all interactions through the platform and blocks attempts to circumvent this process. For example, transaction platforms, such as, Amazon, Booking.com, etc. resist sharing the customers' email addresses with PDEs. Pre- and post-sales interaction between the transaction parties are managed through anonymous alias email addresses. The PDEs thus depend upon the platform to maintain the connection, and, if the PDE loses platform access, then customer access is also lost. To illustrate, YouTubers actively cultivate their community by interacting with their fans to build their followers. For example, when YouTube blocks a creator, they immediately lose access to their fan base and have no way of contacting them to alert them to the new “address”. The motivation to maintain this separation is understandable. To illustrate, eBay uses machine learning to identify violations of its policy forbidding the exchange of contact information between buyers and sellers[8]. Separation from one’s customers effectively ensures “ownership” of customers to the platform and disrupts PDE’s relationship with them.

4.2 Algorithmic management: Ratings, rankings and recommendation systems

Digital platforms utilize algorithmic mechanisms to foster trust between anonymous parties[9], identify reliable vendors, aid in discovery, ensure standards compliance, limit opportunistic behavior, and reduce transaction uncertainty (Tadelis, 2016). In this regard, user-generated ratings and reviews are an essential feature of many platforms because they feed ranking systems that function as screening mechanisms. Review ratings directly influence customer preferences, as Luca (2011) found that a one-star increase in a Yelp rating led to a five to nine percent increase in a restaurant’s revenue and visibility. Ranking systems have become vital, since they enable zero-cost trust creation, monitoring, and a conformity-enforcing mechanism (Ghose, Ipeirotis, & Li, 2014). Effectively, these ranking and review systems shape behavior (Orlikowski & Scott 2012).

Recommendation systems are often vital for platform operations. For example, to help customers discover what they might not find on their own due to the size of the markets, platforms provide personalized recommendation systems. While recommendation systems benefit customers by suggesting products or services tailored to their preferences, they also direct traffic, thus altering market visibility and user action. For PDEs, ranking and recommendation systems are both critical for success and perilous, as they are based on algorithms that can be changed unilaterally.

The algorithms and the data used to regulate rankings, recommendation and discovery are invariably opaque and constantly in flux (Orlikowski & Scott, 2014). For the platform owner, there is little incentive to provide transparency. As a result, PDEs can only speculate on what behavior will satisfy the algorithm. PDEs are thus embedded in a Kafkaesque system, not only of risk, but more seriously, profound uncertainty and vulnerability (Curchod et al. 2019). While algorithmically generated results are often accepted as objective, in fact, they express the platform’s agenda. The algorithms and data upon which they work are opaque, and particularly, the changes in it can appear to be capricious (Scott & Orlikowski, 2012). To illustrate, scores that determine rankings can include a variable that positively values the fact that the ranked firm advertises on the platform. The knowledge that advertising on the platform can affect discovery places great pressure on PDEs to purchase advertising, regardless of whether it provides actual benefits. Consequently, competing PDEs must bid until their profits are reduced to their lowest acceptable level.

4.3 Entry into the PDE's business

Market competition is an intrinsic risk for entrepreneurs. One fundamental risk that a PDE faces, particularly in the case of innovation platforms, is that the platform owner may decide to compete with them. This is particularly potent because, as we mentioned earlier, the owner has a panoptic perspective on all activities (Boudreau & Lakhani, 2009) and the ability to direct traffic towards its offering. The term "asymmetric information access" underappreciates this power (Shapiro & Varian, 1998). The case of Amazon illustrates the use of information to enter a PDE's market space. A former Amazon employee was quoted as saying that Amazon retained "the most valuable data for itself; provides less valuable data to marketplace sellers." The employee continued that the "most valuable info Amazon doesn't share is info about which people have searched for a particular product in the past." Should Amazon decide to enter a particular market niche, it can use this information to "target their private label products with perfect precision" (Capitol Forum, 2018). Although platform entrepreneurs can benefit from valuable information about their products/services (Eckhardt et al., 2018), they only have knowledge about customers that the platform deems beneficial to itself – and the information provided can change as terms and conditions change.

Digital platforms can survey activities on their platform and research market opportunities. With this knowledge, platforms decide whether there are benefits to introducing a targeted competitive product or integrating a specific functionality into the platform's own offerings. For example, Zhu and Liu (2018) showed that Amazon's entry patterns into market segments established by independent merchants targeted entrepreneurs that had high profit margins. This power was described by a former employee:

> Let's say Amazon wants to get into folders. I would find all of the ASINs [Amazon Standard Identification Number] that are being sold on the website now. I'd pull up the history. I'd look at the volumes, price points. Regardless of whether it was sold wholesale or third party, I'd pull it all together. I'd look and see what's the hottest product. What's the hottest variation in color? We'd have these folders in these colors at this price point, and we'd go off and make it ourselves. (Capitol Forum, 2018: 3)

Effectively, in this scenario, the most innovative entrepreneurs operate as scouts for the platform. PDEs innovated new businesses that the platform could then enter and capture, by using its better information and ability to manipulate the platform itself, thereby appropriating the innovator's rents. Alternatively, platforms could decide to raise the fees it charged to successful entrepreneurs to appropriate surpluses. Finally, in the absence of an adequate system of intellectual property protection, such as patents or copyright (Ceccagnoli et al., 2012), the platform can even expropriate the PDEs' businesses. For example, after Microsoft recognizing the potential for Netscape to be a new killer application, it destroyed the new entrant and its business model by bundling Internet Explorer into its operating system (Yoffie & Cusumano, 1998). In effect, Microsoft redesigned the Windows operating system platform to absorb the innovation developed by its ecosystem member, Netscape (Eisenmann et al., 2011).

As a direct consequence of its digital nature, platforms not only broker relationships, but also direct traffic and subsidize the adoption of its offerings, as Amazon has done very effectively. While not always successful in entering a complementor's business, platform owners have a remarkable array of tools to shift the competitive landscape in their favor. In a recent study, Wen and Zhu (2019, p. 16) found that Android app developers responded to Google's threat of market entry and subsequent competition by undertaking "no entry deterrence behavior, such as price reduction and additional innovation . . . [however,] because of the platform owner's power, its entry is unlikely to be deterred". Direct competition from the platform is not simply risk, but a new Knightian uncertainty regarding the defensibility of PDEs' innovations and businesses. Effectively, the Schumpeterian rents "guaranteed" to the successful innovator are at constant risks of appropriation by the platform.

4.4 Changing the terms of participation

For rational actors, market entry is determined through cost-benefit analysis, based upon an understanding of market rules. In a traditional business, the most salient terms of competition are leases; supplier, customer, and competitor relationships; and government regulation. In contrast, a PDE must agree to the platform's terms and conditions for participation (Tiwana, 2014). The key clause in these contracts is that any changes can be made unilaterally.

Changes in the terms of participation regard both the technical components and the rules of engagement. Core issues such as the interface of the platform or the division of revenue between the platform and PDE are decided solely by the platform owner. To illustrate, in fall 2018, eBay unilaterally announced a 12% increase in its commission fees in the Books, DVDs, and Movies categories, while removing the fee discount that eBay Store owners enjoyed[10], thereby directly affecting PDEs' profit margins. For self-published books, Amazon decided that for books priced between

$2.99 and $9.99, the author's share should be 70% of the retail download price. For those priced above or below this range, the share would only be 35%. In this case, authors and publishers' pressure to accept its pricing model, which presumably was the best price for Amazon's goals. Of course, this happens in the non-platform world also, but almost all supplier contracts have fixed terms, whereas the contracts between the platform and PDEs essentially operate as "spot" transactions, in that the terms can be modified at will by the platform.

Entrepreneurs conducting business in a physical store or through their own website are not vulnerable to these shocks. To illustrate, the entrepreneur's landlord cannot, upon seeing their tenant's success, unilaterally abrogate the lease and appropriate the business. Such terms of participation are of critical importance, as they directly affect the emancipatory promise of entrepreneurship (Rindova et al., 2009). In reality, this substantial difference in the terms of participation require surrendering many of the traditional attributes of being an entrepreneur.

4.5 Platform access and delisting

Platforms are private marketplaces and thus access is provided solely at the discretion of the owner. PDEs can be excluded from the platform for undesirable behavior (Evans, 2012), but exclusions can just as easily be "distorted away from pure value creation in the ecosystem towards actions that lead to higher platform profit" (Boudreau & Hagiu, 2009:8). Remarkably, the literature suggests that successful platform owners should be a neutral or, at least, a trusted party (Iansiti & Levien, 2004), perhaps, not recognizing that the owners are for-profit organizations. To illustrate, in return for Apple agreeing to sell on Amazon, the quid pro quo was that the unauthorized independent Apple resellers had their listings removed[11]. In this case, Amazon sacrificed its PDEs for the more valuable Apple account, thereby violating the assumption of neutrality. Paradoxically, the same mechanisms necessary to protect the ecosystem can be used to pursue other goals that advantage the platform.

Exclusion can occur without warning. Additionally, platforms are not required to provide reasons. For a PDE, the decision has immediate repercussions, as their income disappears. Further, the reasons given for suspension are invariably cryptic and platforms provide unclear criteria for adjudicating appeals. Even in the case of a successful appeal, PDEs does not return to status quo ante, as competitors will have displaced them in the rankings. In fact, unethical competitors can report fabricated infractions to the platform[12] (Luca & Zervas, 2016). Effectively, the possibility of delisting means that PDEs' entire business is at constant risk of disappearance. Moreover, the larger and more successful a PDE's business is, the greater the uncertainty and precariousness experienced (Curchod et al. 2019). Moreover, because, many of these platform markets are winner-take-most, there are few alternatives.

5 PDE responses: Power-balancing operations

As power-dependence theory suggests, the subordinate party will try to reduce the power disadvantage in the relationship through what Emerson (1962: 35) termed "balancing operations". Such actions are aimed at altering structural features of the power relationship by reducing the relevance of the resources exchanged and/or by identifying alternative valuable opportunities. Because, in most cases, complete exit is not a viable option due to the winner-take-most aspects of these markets, PDEs have developed responses aimed at mitigating their vulnerability (Kapoor & Agarwal, 2017; Wang & Miller, 2019; Wen & Zhu, 2019). Successfully operationalizing these strategies is difficult because they often challenge a platform's power over the ecosystem (Wen & Zhu, 2019; Wang & Miller, 2019). Thus, a platform's goal is to either stymie or co-opt the strategies discussed below.

5.1 Multihoming

Multihoming refers to a PDE offering a product or service on multiple platforms (Kenney & Pon, 2011), thereby increasing their alternatives (Wang & Miller 2019). There are three general types of multihoming. The first is the classical case, where a PDE operates through multiple platforms (Bresnahan, Orsini, & Yin, 2015). The second type of multihoming is where a PDE uses different channels, e.g., sells on a platform, operates its own website, and may even establish a physical shop (Wang & Miller, 2019). The final type of multihoming is the diversification of income sources discussed in the next section. Often, PDEs combine all three types of multihoming.

5.1.1 Platform multihoming

The costs of multihoming can vary dramatically (Cennamo, Ozalp and Kretschmer, 2018). For example, entrepreneurs selling products on Amazon can easily, with little investment, open a virtual store, listing the same products on the eBay or Etsy platforms. Similarly, for hotels, the costs of multihoming with different online travel agencies are low. In contrast, porting software from iOS to Android or vice versa is more expensive and technically difficult because products must be tailored to platform-specific infrastructure and design (Cennamo et al., 2018). To illustrate, when Snapchat's app update was ported from iPhone to Android, it was buggy, which had a powerful negative impact on revenues[13]. The fact that PDEs must customize their offerings to each platform's specifications is a powerful force for winner-take-most outcomes, as PDEs are unwilling to do so for large numbers of platforms. The decision to multihome is determined by weighing costs against the potential market size (Bresnahan et al. 2015). For example, many PC game firms did not port their games to Apple Macs, as the market was so small that it was not economically justifiable.

Platforms discourage multihoming because it provides PDEs with alternative channels to the market. The tactics used to obstruct multihoming range from designing technological architecture in such a way as to increase the difficulty of multihoming (Cenammo et al. 2018) to prohibiting multihoming in the terms and conditions of the platform's use. Another method is to alter interfaces such as APIs to create incompatibilities, which Apple did to make iOS incompatible with Adobe Flash[14]. To discourage multihoming, platforms make it difficult or impossible for PDEs to inform their audience/customers that they offer the same or similar content on another platform. To illustrate, YouTube terminated the accounts of creators that used their YouTube videos to promote their streams on Twitch, a competitor platform[15]. In certain cases, platforms may recognize the growing power of key PDEs and provide incentives to retain them. Effectively, the possibility of multihoming actualizes a potential threat that PDEs will move their business to another platform.

5.1.2 Channel multihoming

PDEs may also change the balance of power by developing non-platform channels through which to transact. For example, in cases where PDEs and consumers can communicate with one another, it may be possible for them to disintermediate the platform for future transactions. With sufficient trust, PDEs can connect directly with their customers on an off-platform communication medium, thereby excluding the platform and sharing the savings from the platform's fees. Disintermediation is an existential threat not only because it eliminates platform owner's returns, but also because it removes the transactions from the ecosystem (Zhu & Iansiti, 2019). If successful and in sufficient numbers, disintermediation could create an alternative transaction ecosystem.

The two most prominent alternative channels for disintermediation are first, to establish their own website through which customers can purchase goods or services directly and second, to establish a physical store. For both of these strategies to work, a PDE must have the ability to attract traffic. Paradoxically, launching a new channel may include buying online advertising from Google or Facebook platforms to attract customers in the hopes of retaining them for repeated transactions. Another strategy is to use platforms such as the Amazon Marketplace as a marketing platform to connect with potential customers in the hopes that later they can be directed to one's own website. Developing another channel requires greater investment, and it can be initiated only after establishing that there is a market for the product or service outside the platform. Still, it reduces PDEs' dependence upon the platform.

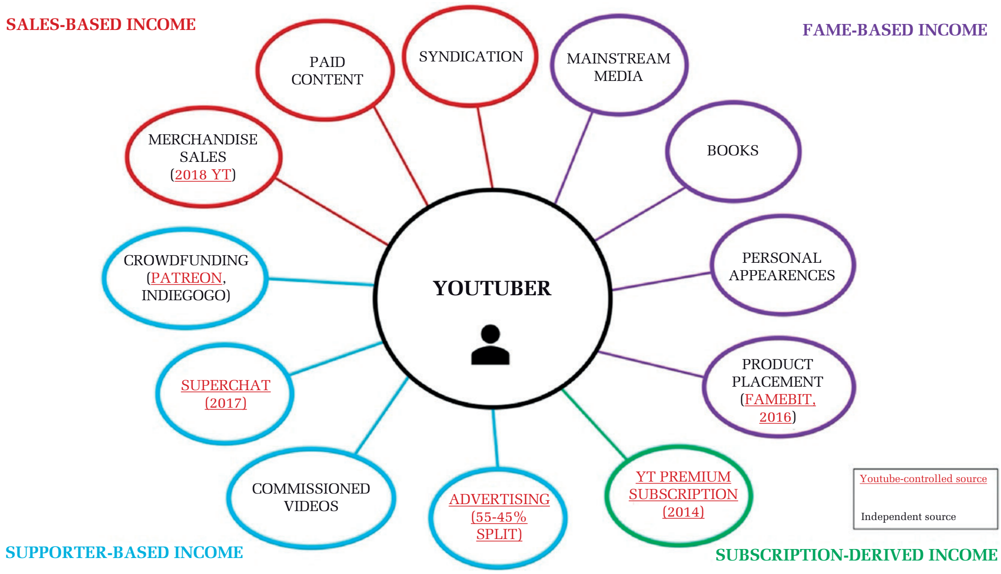
Multi-platform and multi-channel homing provides PDEs with greater market stability and the ability to resist unwelcome changes by the focal platform. Of course, the effectiveness of multiplatform homing can be limited by the lack of alternative platforms. If a PDE is able to attract customers to its own website, then it can increase control and revenue predictability, decrease payments to a platform, and reestablish the ability to interact directly with and learn from one's customers/followers. However, even with an independent website, discoverability continues to be subject to Google Search or influencer recommendations.

5.2 Income diversification

Income diversification is another strategy to mitigate platform power. If extremely successful, off-platform income can increase to the extent that a PDE can leave the platform entirely. On influencer platforms such as YouTube, Instagram, and Pinterest, where a successful PDE can build a large following, they can leverage their fame to create extra-platform income sources. For YouTubers, the direct platform income is a share of advertising revenue. However, because of their strong relationship and direct interaction with users, they can "influence" their audience and generate income from a variety of sources, including but not limited to: personal appearances, merchandise sales, in-video product placements, donations, subscriptions to premium content such as classes, and many other innovative schemes. Income diversification is easier for PDEs in content-based platforms, such as Instagram, Pinterest, or Facebook, since they can grow and leverage their status as public figures. On other platforms, mainly transaction platforms, such as app stores or Etsy, there are far fewer ways to generate an alternative income stream.

Not surprisingly, there can be a tension between PDEs seeking to diversify their income streams and the platforms, which aim to increase their own income and maintain control over their PDEs. Figure 2 summarizes the dialogic evolution of YouTubers' source of income and YouTube attempts to capture either a portion of these alternative income or, at least, to direct it through the platform. For instance, in 2016, to better control sponsorships, YouTube acquired FameBit, a firm that connects creators with brand sponsorships. As part of YouTube, FameBit can provide more granular information about complementors.[16] Now, FameBit has an advantage over competitors and, more importantly, it allows the further "control" of the ability for YouTubers to develop income streams from brand sponsorships. In a similar vein, in 2017, YouTube removed the links YouTubers placed on their channel to direct viewers to their Patreon sites where they could contribute money[17]. As a way of diversifying their income and loosening the hold YouTube had on them, YouTubers contracted to third parties to fulfill their merchandise sales. In response, in 2018 YouTube began a program to introduce "approved" vendors that would fulfill the merchandise sold

Figure 2 YouTubers' income diversification strategies and YouTube ripostes



through channels. This permitted YouTube to gain insight into how much was being sold and to whom[18], and, since YouTube charged the approved vendor a fee, to also increase both its income and its control over the platform-dependent YouTubers.

5.3 Collective action

Collective action is a type of coalition formation (Emerson, 1962) aimed at increasing PDEs' power versus the platform owner. Of course, organizing collective action can be difficult because often, the "public spaces" where PDEs interact are owned by the platform. One mild and unthreatening form is to participate in user forums of various types where PDEs exchange advice and support, and share their experiences (Kuhn & Galloway, 2015). The platforms themselves sponsor these user forums that, unsurprisingly, are not oppositional in nature.

Independent PDE-oriented websites can be venues not only for discussion but also to express grievances. In a number of cases, these have become focal points where resistance to specific changes in platform governance has coalesced. Thus far, most collective action has centered on complaints regarding changes in the terms and conditions[19], and, in certain cases, the platforms have rescinded the changes (Eaton et al., 2015). Such reversals often quell dissent among the PDEs.

There have been cases of more robust collective action, such as the collective withholding of products and services. In November 2018, AbeBooks (owned by Amazon) banned several antiquarian booksellers because their countries did not have acceptable banking institutions for payments. In solidarity with their competitors, hundreds of booksellers removed their listings, and AbeBooks reversed its decision[20]. In this case, the PDEs had alternative market channels and a strong, shared occupational identity that increased solidarity. In July 2019, a German group of YouTubers created a labor organization named "FairTube" and affiliated with IG Metall, the largest German union. Their demands were for YouTube to set up an appeal process that was overseen by a third-party council and provided human contacts for disputes and better explanations about violations, so YouTubers could better understand the decision-making process[21]. While Google agreed to discuss some issues with the organizers of FairTube, it refused to negotiate changes in compensation etc. Collective action can be successful in reversing in changes, though it is more frequent that the platform expresses understanding of the objections by the PDEs but does not reverse the changes.

5.4 Government action

The relationship between a platform and its PDEs is largely within the province of contract law. For this reason, there has been comparatively little litigation by PDEs, as agreed to the terms and conditions when they joined the platform voluntarily and are free to leave. More recently, competition authorities in the European Union have investigated and fined platforms for legal violations. Though Amazon Marketplace has drawn interest from legal scholars and the popular press, most actions, thus far, have not recognized that these vulnerabilities are condition all PDEs experience. Government entry as an additional actor in the PDE-platform relationship could mitigate the dependence of PDEs. To illustrate, small Indian retailers successfully pressured the government to promulgate new rules that make it difficult for retail platforms, such as Amazon and Walmart-owned Flipkart, to sell directly to consumers and operate an online marketplace at the same time[22]. Such actions can prevent a platform from competing directly with its complementors. How far government action will go to change the relationship between platforms and PDEs is uncertain given the current US political debates about platform power.

6 Discussion

Platform ecosystems represent a novel context for entrepreneurship, with peculiar dynamics that contribute to shaping entrepreneurial processes and outcomes. To date, research has mostly focused on platform firms and how they might create an ecosystem and achieve lock-in. Entrepreneurship scholars have principally concentrated and theorized upon how unique affordances of the digital technologies affect entrepreneurship (Nambisan, 2017; von Briel et al., 2018).

In preponderance of the platform and entrepreneurship literature, the relationship between platforms and PDEs as joint ecosystem members is conceptualized in terms of commensalism or mutual benefit. This is a valid but limited perspective. With very few exceptions, commensalism misses a key aspect of this relationship—unequal power between the platform owner and the PDEs. This component should be more clearly recognized in entrepreneurship research on platform-defined markets (see, for example, Nambisan & Baron 2019).

Our analysis is a first step in this direction, as we show how this power imbalance makes platform entrepreneurship substantially different from traditional forms of entrepreneurship. For PDEs, platforms have a contradictory character. New entrants experience a more balanced power-dependence relationship, as the platform owner offers many resources that can allow PDEs to enter the market rapidly and at low cost (Ghazawneh & Henfridsson 2013; Yoffie & Kwak, 2016). In comparison to traditional entrepreneurs, PDEs benefit from a larger population of potential customers, lower entry costs and investment risk in the initial phase of their business. To illustrate, the costs of uploading to YouTube, listing an object on eBay or Amazon, or placing an app in the Apple App Store are trivial. Therefore, new entrants can experiment with part-time activities. In fact, many YouTubers began in their bedroom or dorm room and eBay sellers began by selling miscellaneous items from their home.

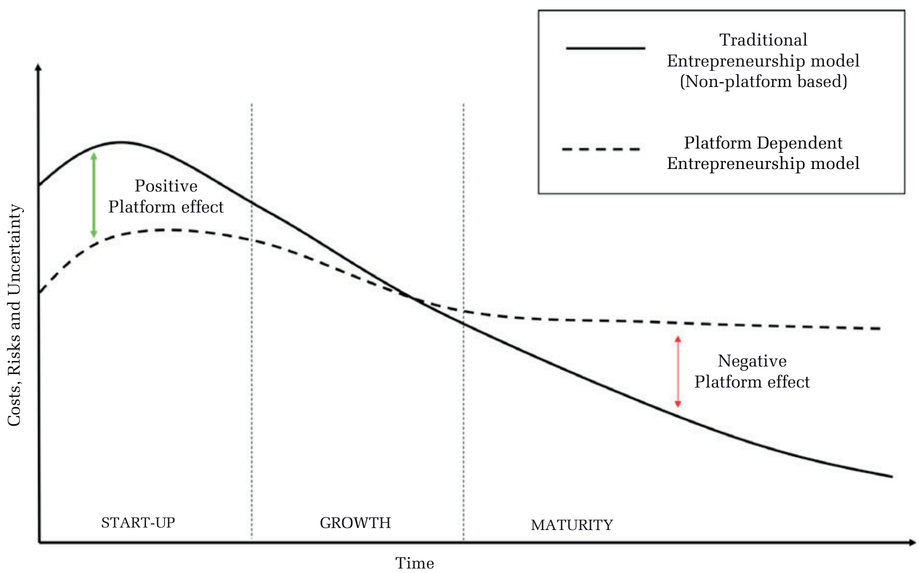
When a PDE's business grows on the platform, the platform's modular infrastructure and boundary resources allow it to scale up at little cost. Platforms may offer PDEs incentives to do so, as successful PDEs create more value for platforms. Whereas a traditional business must invest in infrastructure to meet the demands of growth, a PDE does not need to invest as much because most of the infrastructure is provided by the platform. The platform handles many of the technical and functional issues associated with growth (Ceccagnoli et al., 2012), which is quite important for small businesses, as they are particularly resource-constrained.

As the business matures, growing lock-in effect emerges due to the asset-specificity and lack of portability of the cumulative investment in building their reputation and ranking, transaction history, and ecosystem understanding. In addition, given the winner takes most features of digital platform markets, the PDE's dependence upon the platform increases as the availability of viable business alternatives outside the platform decreases. Mirroring Emerson's argument (1962), as the platform grows and matures, power asymmetries increase while the importance of the individual entrepreneur decreases. Under these circumstances, the incentives provided have a contradictory effect as they reinforce the power imbalance and unique risks emanate from the platform's actions and decisions.

It is possible to illustrate the differences in risk between an entrepreneur establishing an online as opposed to a traditional business. As Figure 3 shows, in comparison to traditional entrepreneurship, platform-dependent entrepreneurship is initially very attractive to entrepreneurs due to the Positive Platform Effect discussed extensively in the literature (Nambisan, 2017; Eckhardt et al., 2018). In the startup phase, the traditional entrepreneur experiences greater difficulty and higher costs of entry because they must secure access to resources and customers. In contrast, PDEs receive these and benefit during their growth phase, as the platform supports scale-up. In contrast, as the traditional entrepreneur's business grows, it owns tangible and intangible assets such as reputation and customers. Thus, it is not as vulnerable to unilateral decisions by another party.

The Negative Platform Effect has received substantially less attention. This effect reflects the novel components of risk that originate from the platform's power over PDEs and strengthens as the platform becomes dominant in a particular market. Entrepreneurship and building a business has always been fraught with risk. Dependence upon a platform, however increases not only risk, but also actually creates a new source of Knightian uncertainty. PDEs' pervasive precarity stems from the fact that this unknowable future distribution of risk extends to the basic tools for doing business, as platform owners can control access to customers, prices, and profit margins—and thereby, the survival of the business. Paradoxically, as the PDE grows,

Figure 3 The different risk profiles of a traditional versus a Platform-Dependent Entrepreneurial firm over Time



while traditional risks may decrease, other things being equal, their dependence and insecurity upon the platform remains—and, in cases of great success, may increase. Successful businesses may in fact become a target for the platform to envelope or for increased usage fees, unless they take actions to build alternatives.

Platform-dependent businesses challenge a number of the key assumptions about competition and value appropriation. For today's entrepreneurs, it is vital to develop a business model that leverages the resources and customers available from a platform to build one's business, while also mitigating the platform's control over that business. PDEs actively implement responses aimed at weakening a platform's grip, by altering the structure of the relationship. Yet, actions meant to counterbalance platform power may be difficult to implement or costly, especially for smaller firms with fewer capabilities (Cennamo et al., 2018).

6.1 Implications for policy-makers

With regard to their ecosystems, platforms are essentially private regulators—a reality that has important policy implications. Not surprisingly, due to their growing and ever more apparent power, platform firms are facing increasing public criticism and regulatory scrutiny. To date, most attention has been directed toward macro-level, anti-competitive dynamics such as the impacts upon public opinion formation, consumers, and data privacy (Furman et al. 2019; Khan, 2016). While all valid concerns, we suggest that governments have not fully grasped how platforms are reshaping the playing field upon which competition and entrepreneurship takes place. Policy-makers have focused less upon the micro-level relationships between platforms and PDEs. Yet, this is where platform power is expressed.

One solution commonly advocated is to dismantle these platforms. Such action should be undertaken with care, as these platforms are also the source of income and livelihoods for an enormous number of entrepreneurs and provide consumers with great variety at attractive prices. Incremental, government regulation could perhaps address the terms and conditions of contracts between PDEs and platforms to ensure that they are not too one-sided. For example, the government could require a reasonable advance notification for any fee and sales commission changes. Delisting or demonetization due to changes in policy should require advance warning and an approved adjudication process. Addressing these terms-and-conditions-related issues could reduce the precarity experienced by PDEs, while not destroying the social benefits that platforms bring.

Policy makers are increasingly considering the implications of platform power. For example, the Indian government recently required Amazon and the Walmart-owned Flipkart to choose between being online retailers and sales platforms to eliminate unfair competition between panoptic platform retailers and PDEs, the competition authorities in the European Union have undertaken a wide variety of actions related to the largest platforms. These include imposing large fines on Google; first, in 2017, for favoring its own shopping website over competitors in its search results, and again, in 2018, for requiring that its apps be pre-installed on Android.

In order to support PDEs to redress power over platforms, in 2019, the European Commission issued a set of rules meant to ensure a "fair, transparent and predictable business environment for smaller businesses and traders on online platforms" (European Commission, 2018, 2019). This EU rule making aimed to alter the terms and conditions set forth in the contracts that PDEs agree to when they join a platform. In 2019, the European Union (2019) enacted a regulation whose explicit goals were to redress the power imbalance by mandating greater transparency and explicit procedures through which the PDEs can file grievances. This initiative and the accompanying regulation, while recognizing the benefits that platforms produce, sets forth a number of requirements regarding the relationship between the platform and the businesses dependent upon it. For example, the 2019 regulation mandates than any major change in the contract between the platform and the PDE requires the provision of 15 days' notice. Platforms are further required to develop transparent dispute-settling mechanisms that businesses selling through a platform can utilize. Moreover, the platform must allow those selling on the platform to have direct contact with their customers, thereby creating an opening to increase direct sales. Finally, the EU regulation requires that platforms provide an account of the main factors used in their online ranking systems and provide advance notice regarding any major changes in such systems. These changes suggest that in Europe the power imbalances that we have identified may be ameliorated, at least, to a certain degree. These European changes also provide openings for future research that can treat the changes as a quasi- experimental setting for understanding the impacts of regulatory changes on PDEs' conditions.

Regulators could promote policies aimed at reducing PDEs' dependence on a single platform. One powerful strategy is to limit platforms' ability to hinder multihoming, thereby increasing competitiveness. If PDEs were more easily able to transfer their businesses to another platform, it would facilitate the entry of new competitive platforms. Currently, most regulatory and antitrust activity is conducted by existing agencies. It might be possible to establish a Platform Competition Authority, whose role would be to investigate PDE complaints and establish a body of regulations aimed at ensuring the viability and health of platform ecosystems.

Governments could also change laws to allow the formation of trade associations or even unions to represent PDEs; something that is currently illegal in US law, as PDEs are treated as independent businesses and not employees. Unfortunately, under current antitrust law because PDEs are businesses this might be seen as forming a cartel and illegally restraining trade. Already, organizations such as the Online Merchants Guild for Amazon merchants and the YouTubers Union formed in Germany for YouTube content creators have emerged[23], but they still lack adequate policy support.

Responses that are more radical could be the formation of stakeholder councils that include the various sides of a platform. For example, councils at YouTube would include representatives of the creators, advertisers, and viewers and councils at Amazon Marketplace would be composed of spokespersons for buyers and vendors. These councils could discuss the implications of major changes on the platform and consider how they would affect ecosystem members. Finally, there has been discussion of forming platform cooperatives, which would create non-profit platforms that operate for the benefit of all stakeholders (Scholz, 2016). An economy, within which platforms are becoming increasingly powerful private regulators, requires the development of novel and innovative regulatory institutions, so that we may continue to reap the benefits of platform-organized markets while ensuring that ecosystem members and the public interest are considered.

6.2 Future Research Directions

The sheer number of PDEs means that entrepreneurship studies must acknowledge their growing relevance in the global economic landscape. Platform-dependent entrepreneurship differs fundamentally from traditional entrepreneurship due to the power asymmetries that define the relationship between PDEs and platform firms. Unraveling the unique risks entrepreneurs face amid powerful platforms lays the groundwork for future research exploring entrepreneurs' platform dependency in greater depth.

We identify several research questions that deserve further attention and we develop a future research agenda around three main area of analysis: the experience of PDEs, the interaction between PDEs and platforms, and the broader implications of this dependence for entrepreneurship in the economy.

6.2.1 The experiences of PDEs

Far more studies are needed on the risks deriving from this power imbalance and on PDEs' strategies to ameliorate them. Competitive dynamics in digital platform markets have peculiar and distinctive features, but managerial scholars have mostly embraced the platform's perspective (Cennamo, 2019; Cusumano et al., 2019). How can PDEs develop and combine strategies to mitigate risks and capture a larger share of the value they create (Wang & Miller, 2019; Wen & Zhu, 2019)? In section 6, we listed many of these strategies, but we do not know which ones are the most effective, under what conditions can an PDE implement them, and, as importantly, the strategies that a platform can use to respond.

Another important research area explores the cost of dependence for PDEs. Further research could develop a more nuanced understanding of how PDEs cope with pervasive uncertainty and the consequent stress, anxiety and precarity that are evident in their accounts of working on a platform (Nambisan & Baron, 2019; Curchod et al., 2019). For example, in exploring the experiences of independent workers in the gig economy, Petriglieri, Ashford & Wrzesniewski (2019) describe how individuals cultivate connections with routines, places, people and broader purposes to deal with the emotional tension of their precarious working conditions. What tactics that make PDEs more resilient?

A related avenue for future studies is the rise of virtual communities of PDEs that offer support and resources to members (Kuhn & Galloway, 2015). Digital entrepreneurship has contributed to an increased distribution of entrepreneurial agency. Investigating how the interplay between competition and cooperation is affected by dependence is a fascinating area for future research.

6.2.2 Interactions between PDEs and platforms

Management of the PDE-platform relationship is essential for the success of both. Platforms offer PDEs easily accessible resources for easy market entry and, if successful, rapid growth. Yet, PDEs face the possibility that a platform firm will change the conditions for success at any moment, potentially without warning or recourse. This means that as entrepreneurs build their business on a platform, they become dependent upon the platform's actions, which are oriented in a balance between sustaining or growing their ecosystem and their own profits.

This paper is a discussion of the general case, and many differences exist due to the remarkable variety of PDEs. Nonetheless, power-dependency is a fundamental constitutive element of the relationship between the platform and its PDEs. Whether PDEs are larger venture capital-financed firms, individuals, and even established businesses transitioning to selling through a platform, their structural position ensures that will experience dependency. Of course, the PDEs' own resources and capabilities affect the degree of dependence (Eaton et al., 2015). For example, in 2017, after a long resistance to selling directly through Amazon, Nike joined the Amazon platform. Because of its market power, Nike was able to negotiate an arrangement with Amazon by which sales from unlicensed Nike distributors and those of knockoff items being sold by third-party sellers would end. In 2019, Nike withdrew from the relationship because it felt that Amazon did not fulfill the agreement and renewed efforts to sell to consumers directly through the Nike website[24]. This example shows how compelling Amazon was, but also the fact that the platform-PDE relationship was so onerous that Nike decided it was better to terminate the relationship. Our work sets the bases for future theorizing on the peculiar features and circumstances that make PDEs more or less dependent upon the platform owner. Identifying the dimensions and their relationships that define the degrees of dependence is a fruitful direction for future entrepreneurship research.

Based on the implications we highlighted, to what extent do entrepreneurs anticipate their dependency when designing business models and strategies for the platform economy and how does this influence their actions? Although digital platforms allow for experimentation with new technologies and business models (Nambisan, 2017; Eckhardt et al., 2018), future research could examine PDEs' degrees of freedom in developing their businesses when platforms can easily identify those creating Schumpeterian rents and attempt to capture those rewards.

It is also important to explore how dependence evolves during the platforms' lifecycle. In order to achieve and maintain a dominant position, platforms owners need to actively and strategically manage the interaction with their complementors over time (McIntyre, Srinivasan, & Chintakananda, 2020), and PDEs will almost certainly face greater demands and higher risks as power asymmetry becomes greater in the later stage of a platform's lifecycle. Rietveld, Ploog, & Nieborgoffer (2020) provide empirical evidence of the increasing costs borne by PDEs when a platform gains dominance, showing that a dominant market position shifts platforms' governance strategy towards profit maximization, and, as a consequence, the value captured by PDEs decreases significantly. The proliferation of user forums where PDEs share and discuss issues (Kuhn & Galloway, 2015) offers a fruitful avenue to further investigate this question empirically, by exploring how PDEs engage with different problems at different stages of a platform's lifecycle.

7 Conclusion

In 2020, platforms are becoming the infrastructure of economy and therefore the context within which entrepreneurship takes place. Launching a new business today requires almost certainly a social media strategy, using online advertising, and deciding whether to offer one's good or service through a platform. We showed that many tenets of traditional notions of entrepreneurship are no longer valid in situations where the entrepreneur depends upon a powerful online digital platform. Paradoxically, this is true despite and because of the fact that the initial investment and risk of establishing a business decreased due to the many resources platforms provide. And yet, building one's business on a platform means facing new dimensions of uncertainty. In particular, platforms are in a powerful position to, in Teece's terminology (2017), "sense" and "seize" the rents that normally accrue to innovators and entrepreneurs. Ultimately, this is because PDEs have no control or little influence over the actions and strategies of platform owners. In fact, in most cases, the can only speculate as to the reasons behind many of changes they experience.

Entrepreneurship education must recognize and incorporate lessons on how entrepreneurs can navigate and manage this new world of affordances and uncertainty. Students must be provided with the knowledge and skills to understand the pitfalls and consequences of their platform-related decisions and have plans to mitigate their dependence. It is vital to increase the awareness among potential entrepreneurs of the paradoxes inherent in building a business on a platform. Platforms are part of the context for entrepreneurship and tools to be used by entreprenerus. To illustrate, an entrepreneur can use Amazon as a marketing forum, while directing repeat consumers to one's own website. Entrepreneurship pedagogy should include case studies to build awareness of alternative strategies, whereby the platform is a resource to achieve the independence that is the promise of entrepreneurship.

We have argued that the fundamental tenets of market capitalism and traditional Schumpeterian notions of entrepreneurship may no longer be valid in the markets where platforms are increasingly powerful. To what extent do we need new theories to understand platform-dependent entrepreneurship? The emergence of platform economy challenges the traditional and iconic vision of entrepreneurial agency, independence, and mastery of one's own destiny. In this environment the "emancipatory potential of entrepreneurship" (Rindova et al., 2009) is threatened by the new risks and multi-dimensional uncertainty that we have chronicled. It is vital for entrepreneurship, strategy, and organizational behavior researchers to better explicate how the platform firms are fundamentally shifting the context for entrepreneurial agency and strategic action.

Endnotes

1 Our definition of platform-dependent entrepreneurs evokes a long-standing debate about who should be considered an entrepreneur (Shane & Venkataraman, 2000). We use the term platform-dependent entrepreneur to indicate individuals or existing organizations entering a platform market. In this regard, platform-dependent entrepreneurship is an inclusive concept that incorporates different entrepreneurial expressions that include app developers on the Apple store, individuals selling on Amazon, but also Instagram influencers and YouTubers. By doing so, we respond to Aldrich & Reuf's (2018) call for a more comprehensive perspective in entrepreneurship research.

[2] For a detailed analysis of users' lock-in in digital businesses see Amit and Zott (2001). The authors focus on the consumers' side, but similar considerations can be made for producers. It is important to note the consumers and producers lock-in is directly linked and further magnified by the presence of indirect network effects (Parker & Van Alstyne, 2005)

[3] ScrapeHero. (2019, April 24). How many products does Amazon sell? – April 2019. Retrieved from https://www.scrapehero.com/number-of-products-on-amazon-april-2019/

[4] Hale, J. (2019, May 7). More than 500 Hours of content are now being uploaded to YouTube every minute. TubeFilter. Retrieved from https://www.tubefilter.com/2019/05/07/number-hours-video-uploaded-to-youtube-per-minute/

[5] Nicas, J. & Collins, K. (2019, September 9) How Apple's apps topped rivals in the App Store it controls. The New York Times. Retrieved from https://www.nytimes.com/interactive/2019/09/09/technology/apple-app-store-competition.html

[6] Alexander, J. (2019, June 25). YouTube looks to demonetization as punishment for major creators, but it doesn't work. The Verge. Retrieved from https://www.theverge.com/2019/6/25/18744246/youtube-

[7] Of course, there is ample anecdotal and other evidence that users or vendors "game" or even violate platform rules (Petre et al. 2019).

[8] Meldner, R. (2017, May 25). eBay messaging now catches more attempts to sell offline. Esellercafe. Retrieved from https://esellercafe.com/ebay-messaging-catches-attempts-to-sell-offline/.

[9] In China, one of the most important elements for Alibaba's success was that it overcame a generalized distrust of strangers, by introducing an escrow system in which the buyer would place the money in escrow with Alibaba that would only release it when the buyer signaled that the transaction was satisfactorily completed (Yu & Shen 2015).

[10] Steiner, I. (2018, August 7). Fall update: eBay raises fees as it makes more demands on sellers. Ecommerce Bytes Blog. Retrieved from https://www.ecommercebytes.com/C/blog/blog.pl?/pl/2018/8/1533654011.html.

[11] Kelley, H. (2018, November 11). What Amazon and Apple's deal means for third-party sellers? CNN. Retrieved from https://www.cnn.com/2018/11/10/tech/amazon-apple-deal-sellers/index.html

[12] Woollacott, E. (2017, September 9). Amazon's fake review problem is now worse than ever, study suggests. Forbes. Retrieved from https://www.forbes.com/sites/emmawoollacott/2017/09/09/exclusive-amazons-fake-review-problem-is-now-worse-than-ever/#1b5c436b7c0f.

[13] Constine, J. (2019, February 5). Snapchat's Android usage keeps falling but rebuild tests well. TechCrunch. Retrieved from https://techcrunch.com/2019/02/05/snapchat-android-rebuild/.

14 Heisler, Y. (2016, December 12). Apple engineer reveals the real reason Steve Jobs didn't allow Flash on the iPhone. BGR https://bgr.com/2016/12/12/steve-jobs-iphone-adobe-flash-testing/

15 Vincent, B. (2018, July 12). YouTube creators are losing their accounts for promoting their Twitch streams. ShackNews. Retrieved from https://www.shacknews.com/article/106108/youtube-creators-are-losing-their-accounts-for-promoting-their-twitch-streams/.

16 Weiss, J. (2018, December 21). FameBit, YouTube's influencer marketing platform, says it can measure organic views like they're ads. TubeFilter. Retrieved from https://www.tubefilter.com/2018/12/21/famebit-measures-organic-views-like-ads/.

17 Kulp, P. (2017, September 28). Bad news for YouTube creators who depend on Patreon. Mashable. Retrieved from https://mashable.com/2017/09/28/youtube-outside-links-partnership-

18 Perez, S. (2018, June 21). YouTube introduces channel memberships, merchandise and premieres. TechCrunch Retrieved from https://techcrunch.com/2018/06/21/youtube-introduces-channel-memberships-merchandise-and-premieres/

19 Dunphy, R. (2017, December 28). Can YouTube survive the adpocalypse? New York Magazine. Retrieved from http://nymag.com/intelligencer/2017/12/can-youtube-survive-the-adpocalypse.html.

20 Flood, A. (2018, November 8). Amazon's AbeBooks backs down after booksellers stage global protest. Guardian. Retrieved from https://www.theguardian.com/books/2018/nov/08/amazon-abebooks-backs-down-after-booksellers-stage-global-protest/.

21 Stephen, B. (2019, August 26) YouTube says it won't negotiate with the YouTubers Union. The Verge. Retrieved from https://www.theverge.com/2019/8/26/20833315/youtube-union-youtubers-negotiate-germany-meeting

22 Phartiyal, (2019, January 31). Walmart, Amazon scrambling to comply with India's new e-commerce rules. Reuters. Retrieved from https://www.reuters.com/article/us-india-ecommerce/walmart-amazon-scrambling-to-comply-with-indias-new-e-commerce-rules-idUSKCN1PP1PN

23 Webb, A. (2019, August 2013). Watch out Google, YouTubers are unionizing. Bloomberg. Retrieved from https://www.bloomberg.com/opinion/articles/2019-08-14/german-union-ig-metall-is-backing-youtubers-fighting-google

24 Porter, J. (2019, November 19). "Nike ends direct sale of clothes and shoes on Amazon." *The Verge*. Retrieved from https://www.theverge.com/2019/11/14/20964428/nike-amazon-wholesale-deal-clothes-shoes-third-party-sellers-gray-market-counterfeit

References

Adner, R. (2017). Ecosystem as structure: An actionable construct for strategy. *Journal of Management*, *43*(1), 39-58.

Aldrich, H. E., & Ruef, M. (2018). Unicorns, gazelles, and other distractions on the way to understanding real entrepreneurship in the United States. *Academy of Management Perspectives*, *32*(4), 458-472.

Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic Management Journal*, 22(6-7), 493-520.

Arthur, B. (1989). Competing technologies, increasing returns, and lock-in by historical events. *Economic Journal*, 99(394), pp. 116-131.

Autio, E., Kenney, M., Mustar, P., Siegel, D., & Wright, M. (2014). Entrepreneurial innovation: The importance of context. *Research Policy, 43*(7), 1097–1108.

Autio, E., Nambisan, S., Thomas, L. D., & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. *Strategic Entrepreneurship Journal*, *12*(1), 72-95.

Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In Gawer, A. (Ed.), *Platforms, markets and innovation* (pp. 19-44). Cheltenham, UK: Edward Elgar.

Boudreau, K. J., & Hagiu, A. (2009). Platform rules: Multi-sided platforms as regulators. In Gawer, A. (Ed.), *Platforms, Markets and Innovation* (pp. 163-191). Cheltenham, UK: Edward Elgar.

Boudreau, K. J., & Jeppesen, L. B. (2015). Unpaid crowd complementors: The platform network effect mirage. *Strategic Management Journal*, *36*(12), 1761-1777.

Boudreau, K., & Lakhani, K. (2009). How to manage outside innovation. *MIT Sloan Management Review*, *50*(4), 69-76.

Bresnahan T, Orsini J, & Yin P (2015) Demand heterogeneity, inframarginal multihoming, and platform market stability: Mobile app. Working paper, Stanford University, Stanford.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: W.W. Norton.

Capitol Forum. (2018). Amazon: EC investigation to focus on whether Amazon uses data to develop and favor private label products. *Capitol Forum 6*(393), 1-4. Retrieved from https://thecapitolforum.com/wp-content/uploads/2018/11/Amazon-2018.11.05.pdf

Ceccagnoli, M., Forman, C., Huang, P., & Wu, D. J. (2012). Co-creation of Value in a Platform Ecosystem! The Case of Enterprise Software. *MIS Quarterly*, *36*(1), 263.

Cennamo, C., & Santalo, J. (2013). Platform competition: Strategic trade-offs in platform markets. *Strategic Management Journal*, *34*(11), 1331-1350.

Cennamo, C., Ozalp, H., & Kretschmer, T. (2018). Platform architecture and quality trade-offs of multihoming complements. *Information Systems Research*, *29*(2), 461-478.

Cennamo, C. (2019). Competing in digital markets: A platform-based perspective. *Academy of Management Perspectives*. In press.

Cook, K. S., Emerson, R. M., Gillmore, M. R., & Yamagishi, T. (1983). The distribution of power in exchange networks: Theory and experimental results. *American Journal of Sociology*, 89(2), 275-305.

Curchod, C., Patriotta, G., Cohen, L., & Neysen, N. (2019). Working for an algorithm: Power asymmetries and agency in online work settings. *Administrative Science Quarterly*, Available at http://dx.doi.org/ 10.1177/0001839219867024.

Cusumano, M. A., Gawer, A., & Yoffie, D. B. (2019). *The Business of Platforms: Strategy in the Age of Digital Competition, Innovation, and Power*. New York: HarperCollins Publishers.

Eaton, B., Elalouf-Calderwood, S., Sorensen, C., & Yoo, Y. (2015). Distributed tuning of boundary resources: The case of Apple's iOS service system. *MIS Quarterly*, 39(1), 217-243.

Eckhardt, J. T., Ciuchta, M. P., & Carpenter, M. (2018). Open innovation, information, and entrepreneurship within platform ecosystems. *Strategic Entrepreneurship Journal*, *12*(3), 369-391.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they? *Strategic Management Journal*, *21*(10-11), 1105-1121.

Eisenmann, T., Parker, G., & Van Alstyne, M. (2011). Platform envelopment. *Strategic Management Journal*, *32*(12), 1270–1285.

Emerson, R. M. (1962). Power-dependence relations. *American Sociological Review*, 27(1), 31-41.

European Union. (2019). Regulation (EU) 2019/1150 of the European Parliament and the Council of the European Union. (June 20) https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1568719209879&uri=CELEX:32019R1150

European Commission. (2019, February). Digital single market: EU negotiators agree to set up new European rules to improve fairness of online platforms' trading practices. Press Release. Retrieved from http://europa.eu/rapid/press-release_IP-19-1168_en.htm.

European Commission. (2018). Commission Staff Working document - Impact assessment and Executive Summary accompanying the document Proposal for a Regulation of the European Parliament and of the Council on promoting fairness and transparency for business users of online intermediation services. https://ec.europa.eu/digital-single-market/en/news/impact-assessment-proposal--promoting-fairness-transparency-online-platforms

Evans, D. S., Hagiu, A., & Schmalensee, R. (2006). *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries.* Cambridge: MIT Press.

Evans, D. S. (2012). Governing bad behavior by users of multi-sided platforms. *Berkeley Technology Law Journal, 27*, 1201-1250.

Evans, D. S., & Schmalensee, R. (2016). *Matchmakers: The New Economics of Multi-sided Platforms*. Harvard Business Review Press.

Furman, J., Coyle, D., Fletcher, A., McAules, D., & Marsden, P. (2019). *Unlocking Digital Competition: Report of the Digital Competition Expert Panel*. HM Treasury, United Kingdom.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, *43*(7), 1239-1249.

Gawer, A., & Cusumano, M. A. (2002). *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation*. Boston: Harvard Business School Press.

Gawer, A., & Cusumano, M. A. (2014). Industry platforms and ecosystem innovation: Platforms and innovation. *Journal of Product Innovation Management*, *31*(3), 417–433.

Gereffi, G., Humphrey, J., & Sturgeon, T. (2005). The governance of global value chains. *Review of International Political Economy*, 12(1), 78-104.

Gerwe, O., & Silva, R. (2018). Clarifying the sharing economy: Conceptualization, typology, antecedents, and effects. *Academy of Management Perspectives*, *34*(1), 65-96.

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

Ghose, A., Ipeirotis, P. G., & Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science*, *60*(7), 1632-1654.

Iansiti, M., & Levien, R. (2004). Strategy as ecology. *Harvard Business Review, 82*(3), 68-81.

Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, 39(8), 2255-2276.

Kapoor, R., & Agarwal, S. (2017). Sustaining superior performance in business ecosystems: Evidence from application software developers in the iOS and Android smartphone ecosystems. *Organization Science*, *28*(3), 531-551.

Katila, R., Rosenberger, J. D., & Eisenhardt, K. M. (2008). Swimming with sharks: Technology ventures, defense mechanisms and corporate relationships. *Administrative Science Quarterly*, 53(2), 295-332.

Kenney, M. & Pon, B. (2011). Structuring the smartphone industry: Is the mobile internet OS platform the key? *Journal of Industry, Competition and Trade, 11*(3), 239-261.

Kenney, M., & Zysman, J. (2016). The rise of the platform economy. *Issues in Science and Technology*, *32*(3), 61-69.

Khan, L. M. (2016). Amazon's antitrust paradox. *Yale Law Journal*, *126*, 710-805.

Knight, F. (1921). *Risk, Uncertainty and Profit*. New York: Augustus Kelley.

Kuhn, K. M., & Galloway, T. L. (2015). With a little help from my competitors: Peer networking among artisan entrepreneurs. *Entrepreneurship: Theory and Practice*, *39*(3), 571-600.

Luca, M. (2011). Reviews, reputation, and revenue: The case of Yelp.com. Working Paper 12-016, Harvard Business School, Boston.

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, *62*(12), 3412-3427.

McIntyre, D. P., & Srinivasan, A. (2017). Networks, platforms, and strategy: Emerging views and next steps. *Strategic Management Journal*, *38*(1), 141-160.

McIntyre, D. P., Srinivasan, A., & Chintakananda, A. (2020). The persistence of platforms: The role of network, platform, and complementor attributes. *Long Range Planning*, 101987.

Miric, M., Boudreau, K. J., & Jeppesen, L. B. (2019). Protecting their digital assets: the use of formal & informal appropriability strategies by app developers. *Research Policy*, 1-13.

Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. *Entrepreneurship Theory and Practice*, *41*(6), 1029–1055.

Nambisan, S., & Baron, R. A. (2013). Entrepreneurship in innovation ecosystems: Entrepreneurs' self–regulatory processes and their implications for new venture success. *Entrepreneurship: Theory and Practice*, *37*(5), 1071-1097.

Nambisan, S., & Baron, R. A. (2019). On the costs of digital entrepreneurship: Role conflict, stress, and venture performance in digital platform-based ecosystems. *Journal of Business Research*.

Nambisan, S., Siegel, D., & Kenney, M. (2018). On open innovation, platforms, and entrepreneurship. *Strategic Entrepreneurship Journal*, *12*(3), 354-368.

Orlikowski, W. J. & Scott, S.V. (2014). What happens when evaluation goes online? Exploring apparatuses of valuation in the travel sector. *Organization Science*, 25 (3), 868-891.

Ozalp, H., Cennamo, C., & Gawer, A. (2018). Disruption in platform-based ecosystems. *Journal of Management Studies*, 55(7), 1203-1241.

Parker, G. G., & Van Alstyne, M. W. (2005). Two-sided network effects: A theory of information product design. *Management Science*, 51(10), 1494-1504.

Parker, G., Van Alstyne, M., & Choudary, S. P. (2016). *Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You*. New York: W.W. Norton.

Petre, C., Duffy, B. E., & Hund, E. (2019). "Gaming the system": Platform paternalism and the politics of algorithmic visibility. *Social Media+ Society*, 5(4), 1-12.

Petriglieri, G., Ashford, S. J., & Wrzesniewski, A. (2019). Agony and ecstasy in the gig economy: Cultivating holding environments for precarious and personalized work identities. *Administrative Science Quarterly*, 64(1), 124-170.

Rietveld, J., Ploog, J. N., & Nieborg, D. B. (2020). The Coevolution of Platform Dominance and Governance Strategies: Effects on Complementor Performance Outcomes. *Academy of Management Discoveries*, In press.

Rindova, V., Barry, D. & Ketchen, D. (2009). Entrepreneuring as emancipation. *Academy of Management Review, 34*(3), 477-491.

Rosenblat, A. (2018). *Uberland: How Algorithms Are Rewriting the Rules of Work*. Berkeley: University of California Press.

Schilling, M. A. (2002). Technology success and failure in winner-take-all markets: The impact of learning orientation, timing, and network externalities. *Academy of Management Journal*, *45*(2), 387-398.

Scholz, T. (2016). *Platform Cooperativism: Challenging the Corporate Sharing Economy*. New York: Rosa Luxemburg Foundation.

Schumpeter, J. A. (1942). *Capitalism, Socialism, and Democracy*. New York: Harper.

Search Engine Journal. (2019). History of Google algorithm updates. Retrieved from https://www.searchenginejournal.com/google-algorithm-history/

Shane, S., & Venkataraman, S. (2000). The promise of entrepreneurship as a field of research. *The Academy of Management Review*, *25*(1), 217.

Shapiro, C. & Varian, H. R. (1998). *Information Rules: A Strategic Guide to the Network Economy*. Boston: Harvard Business Press.

Tadelis, S. (2016). Reputation and feedback systems in online platform markets. *Annual Review of Economics, 8*(1), 321-340.

Täuscher, K., & Kietzmann, J. (2017). Learning from failures in the sharing economy. *MIS Quarterly Executive*, *16*(4), 253–264.

Teece, D. J. (2017). Dynamic capabilities and (digital) platform lifecycles. In *Entrepreneurship, Innovation, and Platforms* (211-225). Emerald Publishing Limited.

Thies, F., Wessel, M., & Benlian, A. (2018). Network effects on crowdfunding platforms: Exploring the implications of relaxing input control. *Information Systems Journal, 28*(6), 1239–1262.

Tiwana, A. (2014). *Platform Ecosystems: Aligning Architecture, Governance, and Strategy*. Waltham, MA: Morgan Kaufmann.

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

von Briel, F., Davidsson, P., & Recker, J. (2018). Digital technologies as external enablers of new venture creation in the IT hardware sector. *Entrepreneurship: Theory and Practice*, *42*(1), 47–69.

von Hippel, E. 1988. *Sources of Innovation*. Oxford: Oxford University Press.

Wang, R. D., & Miller, C. D. (2019) Complementors' engagement in an ecosystem: A study of e-book offerings on Amazon Kindle. *Strategic Management Journal, 41*(1), 3-26.

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

Wen, W., & Zhu, F. (2019). Threat of platform-owner entry and complementor responses: Evidence from the mobile app market. *Strategic Management Journal*, 40(9), 1336–1367

Yamin, M., Sinkovics, R. R., Lee, J., & Gereffi, G. (2015). Global value chains, rising power firms and economic and social upgrading. *Critical Perspectives on International Business*, 11(1), 319-339.

Yoffie, D. B., & Cusumano, M. A. (1998). *Competing on Internet Time: Lessons from Netscape and Its Battle with Microsoft*. New York: Simon and Schuster.

Yoffie, D. B., & Kwak, M. (2006). With friends like these: The art of managing complementors. *Harvard Business Review*, *84*(9), 88-98.

Zhu, F., & Liu, Q. (2018). Competing with complementors: An empirical look at Amazon.com. *Strategic Management Journal, 39*(10), 2618-2642.

Zhu, F., & Iansiti, M. (2019). Why some platforms thrive and others don't. *Harvard Business Review, 97*(1), 118–125.

Zittrain, J. (2008). *The Future of the Internet--and How to Stop It*. New Haven: Yale University Press.

Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Profile Books.

ARTICLE 8

Supranationalism, Sino-American Technology Separation, and Semiconductors:

First Observations

Harald Holmström[a]
Timo Seppälä[a, b]

Affiliations:

[a] Aalto University, Finland
[b] ETLA Economic Research, Finland

Original publication:

Holmström Harald and Seppälä Timo. (10.9.2020). "Supranationalism, Sino-American Technology Separation, and Semiconductors: First Observations". ETLA Working Papers No 82.

Abstract

Global forces have shaped the world since the industrial and digital ages. A recent perspective on globalization acknowledges the growth of three supranational economic, social, and manufacturing blocs, namely the USA, the EU, and most recently, China. In this larger picture China contends with the US to become the largest economy in the world. Recent developments in the US–China trade conflict have centered on digital technology and have set the two countries on a path towards a technology separation. This technology separation will disrupt the unique and strategically important global value chains of digital technologies. We define digital technologies as the stack of integrated hardware and software systems that enable various end applications to emerge from computation.

The technology separation will happen in the lower hardware levels of the technology stack, that is, in knowledge- and capital-intensive semiconductor technology, design and manufacturing. A separation within semiconductor technology could have serious implications for Europe, but especially for smaller open economies such as that of Finland. The key to designing Europe's semiconductor technology strategy is understanding the history, technologies, and dynamics of the semiconductor industry as well as understanding industrial policies regarding semiconductors in the USA and China. What are the different options for Europe if the technological separation continues?

Keywords

Semiconductors, Semiconductor industry, Digital technology, Technology stack, Sino-American technology separation, Industrial policy

1 Tectonic shifts in the global world order

Different forces of globalization have shaped the world since the industrial and digital ages.[1] Additionally, globalization has made nations more integrated and interdependent through diverse networks of cross-border relationships (Baldwin, 2006). Most large multinational corporations (MNCs) trade regionally based on nationally located assets, and financial flows have been concentrated to North America, Europe, and East Asia (Seppälä, Kenney, & Ali-Yrkkö, 2014).

The contemporary view of globalization is based on recent events and acknowledges the progress of supranational economic, social and manufacturing blocks, namely the USA, the EU, and most recently, China (Seppälä, Kenney, & Ali-Yrkkö, 2014; Hirst, Thompson, & Bromley, 2015). The relationships between these economic, social and manufacturing blocs, and their overlapping interests govern global trade, industry, digitization, and technologies. Additionally, differing ideologies and modi operandi undermine multilateral endeavors. From this perspective, it has become evident that China is contending with the US to become the largest economy in the world (*The Economist*, 2020a; Frankel, 2020).

The Chinese state has assumed a large role in providing support for industrialization (Nolan, 2001; Harrison, 2014). Wade (1990) posited that late industrializers typically go through a distinct phase of state intervention and protectionism in order to develop domestic industries. It is also widely known that industrial policy and government intervention aimed at building technological competence have served as the driving forces behind late industrialization in advanced electronics industries, for instance, in Japan and South Korea (Amsden, 1989; Wade, 1990).

China initially entered labor-intensive parts of electronics value chains in the 1990s and later, those of semiconductor value chains in the 2000s, mainly through Taiwanese and American foreign direct investment (FDI) (Brown & Linden, 2005). This FDI made China the largest exporter of computers around 2004 (Yang, 2006). In the beginning of China's upgrading journey, as much as 90% of value adding components had to be imported from other nations (Assche & Gangnes, 2010).[2] It has later been documented that China has captured a larger share of value creation in the electronics supply chains (Larsen, Seppälä & Ali-Yrkkö, 2018). Furthermore, the Chinese state continues to provide strong support for its domestic technology industries (see the Made in China 2025 initiative [Zenglein & Holzmann, 2019]).

Recent developments in the US–China trade conflict have centered on digital technology and have set the two countries on a path towards technological separation (*The Economist*, 2020b). The US invokes national security concerns over 5G networks and it has targeted Huawei, the Chinese exporter of telecom network equipment and smartphones. To concretize, Huawei was first added to the Department of Commerce entity list in May 2019, requiring export licenses for American firms to continue supplying Huawei (Department of Commerce, 2019). Further Huawei export restrictions on integrated circuits (ICs) produced using American equipment were announced in May 2020 (Department of Commerce, 2020). The latest trade restrictions in the semiconductor value chain are particularly interesting as they affect China indirectly through Taiwan Semiconductor Manufacturing Company (TSMC).[3]

When it comes to the hardware (HW) and software (SW) digital technology stack, China has demonstrated its competitiveness in digital platforms (e.g., TikTok, WeChat and Alibaba) and digital systems (e.g., Huawei, Xiaomi, Oppo). Yet the country lacks self-sufficiency in semiconductors—the lower hardware layers of the technology stack. Discussions, policies, and actions relating to digital platforms and systems will accordingly have significance but arguably not be as important and decisive as those regarding semiconductors.

Semiconductors are essential to modern life. New digital technologies—such as edge computing, industry 4.0, general artificial intelligence (AI), and quantum computing—rely heavily on semiconductor progress in delivering their promise of massive benefits to the global economy. Leading-edge semiconductors are also seen as "critical to defense systems and US military strength" (PCAST, 2017). Additionally, the global and distributed nature of IC value chains pose hardware security risks, and ensuring the integrity of ubiquitous semiconductor devices is hence important in order to mitigate future cybersecurity threats (Rostami, Koushanfar, & Karri, 2014).

Computation with semiconductors has become a cornerstone of scientific research and the human ability to solve increasingly complex problems relies on digital technologies, that is, on "the synergy of advanced algorithms, data and hardware" (PRACE, 2018). It is quite trivial, then, to see that quicker and more pervasive computation with greater power efficiency can benefit the public and equally provide a strategic edge in national security and business.

The motivation for writing this working paper is a concern that Europe, including Finland, will fall behind China and the US in the development of the digital technologies that will drive economic growth in the future as the technology separation continues. Furthermore, Europe and Finland need to reconsider their technology strategies if the separation affects the semiconductor layers of the digital technology stack (see Figure 1 on the next page). What options does Europe have to secure technological sovereignty[4] if the Sino-American technology separation continues? Does Europe need to achieve technological sovereignty in semiconductors?[5]

The current European-wide technology strategy envisions developing a high-quality digital infrastructure by increasing EU, member state, and industry technology investments[6] to €20 billion annually in order to keep up with the US and Asia (European Commission, 2020). The goal is to secure "technological sovereignty in key enabling technologies and infrastructure" (European Commission, 2020). While there has been widespread discussion on American platform giants' market power in Europe, we want to bring semiconductor technology into European policy discussions.

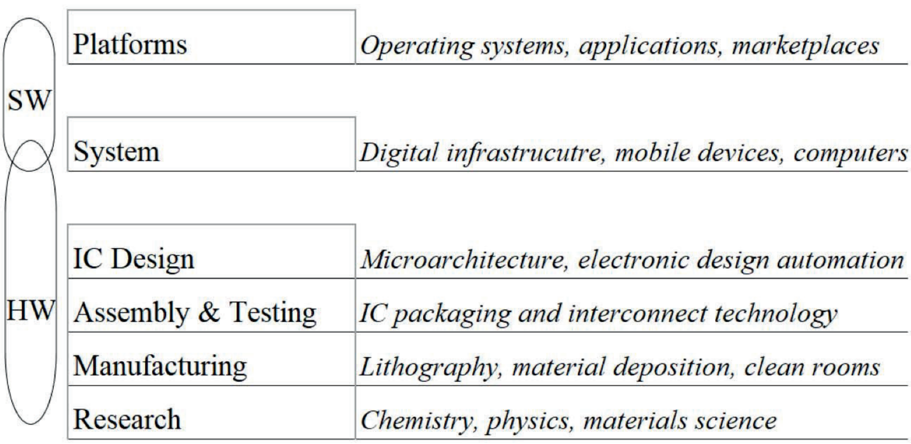
The European Commission's AI white paper is an attempt to steer technology industry development in Europe. However, Europe's strategy with regard to semiconductors remains unclear. The commission states that initiatives such as the European Processor Initiative (EPI) might reduce the dominance of non-EU players in the semiconductor markets (European Commission, 2020). However, meaningful achievements in upgrading the European semiconductor industry remain unlikely with the current strategy and current levels of investment.

This working paper continues as follows: First, we provide a definition of the semiconductor technology stack. Second, we write about the historical context of global value chains in the semiconductor industry. Third, we describe the current state of semiconductor manufacturing and how value is captured and created geographically. Fourth, we present how globally significant supranational blocs have acted in support and the policies of their semiconductor industries. We conclude that political action on European technology sovereignty might pose a threat to Finnish technology neutrality and respective exports in the future.

2 Defining the semiconductor technology stack

We define *digital technologies* as the integrated hardware and software systems that enable (and have enabled) various end applications to emerge from computation. The technology stack has been used to describe strategies and dynamics in the electronics industry and mobile internet (see, e.g., Kenney & Pon, 2012). We adopt a deeper view of the stack in order to capture how semiconductors affect the global technology competition. The hardware and software layers are depicted in Figure 1 below.

Figure 1 The hardware and software stack of digital technologies



The top layer of the technology stack is *platforms*, an umbrella term that we use for operating systems, applications, marketplaces, and social networks. The platform layer remains largely unaffected by the trade war because only 3% of US software industry revenue was generated in China in 2019 (*The Economist*, 2020b) and vice versa. Furthermore, it can be noted that the open source standards, application programing interfaces, and easy reproducibility of software reduce the significance of software in the conflict. American platform companies that allegedly have large market power have been scrutinized, especially in Europe, while China has managed to cultivate its own breed of domestic technology giants. European consumers are currently free to choose between American and Chinese platforms.

In our view, the system layer bridges the hardware and software domains. It provides a category for a diverse range of companies providing telecommunications infrastructure, mobile devices, and computers. In this layer the functionality, performance, connectivity, and security (among other attributes) of digital technologies are defined. Without systems companies, there would not be any smartphones or computers, nor any wireless networks for platforms to provide their offerings. As opposed to the US, China and Europe are self-sufficient in telecommunications networks (Huawei, Nokia, & Ericsson). Global value chains in the electronics industry are heavily reliant on China, with China being the largest exporter of electronics.

The lower hardware layers have become a flashpoint in the Sino-American trade conflict. Semiconductor ICs are the foundation for computation in data centers, smartphones, PCs, aerospace, business, national defense, and healthcare. They underpin the estimated revenue of $2 trillion in global e-commerce (*The Economist*, 2018), and national leadership in semiconductors is strategic. The semiconductor industry enables both the system and platform levels of the digital technology stack. The US is trying to maintain its technology leadership by restricting Chinese access to leading-edge ICs while simultaneously accelerating innovation efforts at home. It is interesting to note that platform companies have begun moving down the stack by investing in proprietary chip designs to accelerate workloads in their computing environments (e.g., Google TPU [Cherney, 2020], Alibaba Hanguang [Kharpal, 2019]). There is however a clear distinction between the design and manufacturing of ICs—Apple is for instance making its own semiconductor designs but relies on TSMC for manufacturing.

The current positions of the supranational economic and social blocs in the semiconductor technology stack are indicated in Table 1 on the next page.

3 Semiconductors—a flashpoint in the US–China trade war

3.1 The US–Japan trade war in the 1980s

The innovation, competitiveness, and integrity of the US semiconductor industry is now facing challenges (PCAST, 2017). However, the prospect of a US deterioration in semiconductors because of foreign competition is nothing new. US semiconductor companies faced intense competition from Japanese dynamic random-access memory (DRAM) manufacturers in the 1980s (Brown & Linden, 2011). It took about 20 years for Japanese manufacturers to achieve technological parity with the US: in the 1960s, government agencies forced technology transfers from foreign companies (e.g., IBM) seeking access to the Japanese market (Prestowitz, 1988). The Japanese government furthermore actively subsidized research and promoted cooperation between its intensely competing business groups (Fransman, 1990).

By developing superior manufacturing capabilities, Japanese semiconductor divisions surpassed the US in both market share and R&D expenditure (Brown & Linden, 2011). Crashing demand for DRAM in 1985 and eager Japanese investments led to severe overcapacity and an acute crisis in US semiconductor manufacturing. US firms weathered the storm through consolidation and repositioning from DRAM toward custom logic processors. Industry collaboration simultaneously increased as the Semiconductor Industry Association (SIA) was formed to solicit support from the government amid calls for "fair trade" and the elimination of "dumping" in US and third markets, although the latter was never proven (Irwin, 1996).

Despite the rhetoric of semiconductors being strategic high technology, Irwin (1996) concluded that the 1980s DRAM dispute followed a similar pattern to that of other instances of trade friction. Namely, that the rapid entry of reasonably priced high-quality Japanese goods (e.g., cars and textiles) was a shock to isolated American manufacturers. The resolution of the US–Japan rivalry adopted numerical targets, so-called managed trade (see Flamm, 2010), for US access to the Japanese market (Irwin, 1996). In an interesting precedent to strategic high technology, US trade policy shifted away from setting trade "rules" and moved towards seeking a transaction-

Table 1 The current position of each economic and social bloc in the semiconductor technology stack

	USA	China	Europe
Platforms	A leading role	Local platform firms since 2000s	A minor role
Systems	Equipotent status		
Hardware	A leading role	On-going upgrading since 2000s	A minor role

al "outcome" (Dick, 1996). The threat of US sanctions reduced the scope for direct government intervention in established industries (Brown & Linden, 2011). But, because Japanese manufacturers could sell to Europe and easily circumvent voluntary export restrictions, some argue that the extensive integration of semiconductor markets rendered the US unilateral approach inefficient in the short term (Dick, 1996).

Although a policy response might not work exactly as intended, history shows that industry leaders can collaborate and consolidate in order to lobby for support when facing an exogenous crisis. The current challenge to the US semiconductor industry however has a different nature. China plays a dual role in the ongoing conflict as it is developing its domestic semiconductor capabilities while simultaneously guarding the largest and fastest growing market for semiconductors globally. While the US sought to manage its trade deficit with Japan, the current goal of the US government is decoupling from China (Koskinen, 2020).

3.2 A brief history of TSMC

The nurturing and flourishing of Taiwan's semiconductor industry form one of the most successful cases of industrial establishment. The two main influences on Taiwan's success in the semiconductor industry are detailed in the related literature. The first was the institutional view of an innovative public-private partnership that enabled the diffusion of technologies and knowledge to private firms (Mathews, 1997). The second was highlighting the role of engineers and scientists with US educational and professional experience returning to Taiwan (so-called returnees; Saxenian, 2006), although these returnees mainly participated in later industry development (Kenney, Breznitz & Murphree, 2013). TSMC, for instance, benefited from returnees by recruiting many of them to senior management positions, which provided vital business connections in addition to managerial capabilities (Saxenian, 2006). Progressive integration into formal corporate production networks and informal knowledge networks helped Taiwan upgrade its technical capabilities and thus sustained its semiconductor industry's competitiveness (Ernst, 2010).

TSMC is the technology leader in semiconductor fabrication and can be seen as a bottleneck in the semiconductor value chains from the American perspective. At the height of the US–Japan DRAM crisis, TSMC was spun off from a pilot project within the Electronics Research Service Organization (ERSO) in 1985. ERSO had made several technology transfers from various actors in order to expand its technical semiconductor capabilities. Taiwan's first semiconductor fabricator, United Microelectronics Corporation (UMC), was created as an ERSO technology and staff spin-off with government financing in 1980. While taking over ERSO's manufacturing pilot, the new company, TSMC, was formed as a joint venture with Dutch multinational Philips. In return for an advantageous position in Taiwan's semiconductor industry, Philips transferred both its existing technology (which trailed the world leading-edge by 1–2 generations) and its cross-licensing agreements with other manufacturers to the new joint venture. The last detail effectively shielded TSMC from intellectual property (IP) rights disputes that plagued other East Asian manufacturers (Mathews, 1997).

By the mid-1990s, TSMC had retained its cost advantage while achieving technological parity with leading IDMs and foundries in the United States and Japan (Saxenian, 2006). All in all, the Taiwanese upgrade to the leading edge took 20 years (fundamental capabilities were being nurtured 10 years prior to TSMC's entry).

TSMC's success is founded in its reliability in regard to delivering timely manufacturing process advancements. The company pioneered the innovative pure-play foundry business model when it was conceived in the 1980s by TSMC's founder, Dr. Morris Chang. Chang had worked at Texas Instruments for 25 years and noticed a trend of top engineers founding their own semiconductor businesses. But these start-ups could not finance huge capital expenditure in semiconductor manufacturing, and Chang thus identified a new market opportunity (Nenni, 2020).

The fragmentation of the semiconductor industry started in the 1970s when integrated device manufacturers (IDMs; such as Intel), along with independent equipment and materials producers, were founded in Silicon Valley. TSMC capitalized on the beginning fragmentation by focusing purely on the manufacturing process and catering to a newly established chip design industry. Through a design–manufacturing partnership, semiconductor foundries benefited from having access to developing (novel) end markets and the design firms gained access to leading-edge manufacturing without the huge capital commitments required for a fab. The availability of electronic design automation (EDA) tools and standardization through IP blocks facilitated the entry of design firms without manufacturing capabilities. IC manufacturing was further unbundled by third firms specializing in the final assembly and testing of ICs. Fragmentation has driven innovation and allowed specialized firms throughout the value chain to generate value with innovation in new products, materials, microarchitectures, manufacturing processes, and IC packaging (Saxenian, 2006).

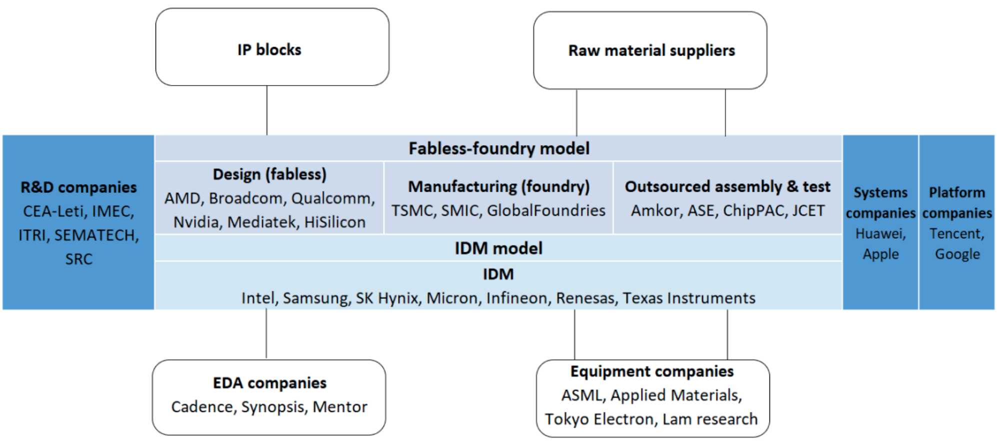
3.3 The current semiconductor value chain

Manufacturing ICs from silicon requires one of the most complex manufacturing processes on earth, and the semiconductor industry constitutes a great but idiosyncratic example of a global value chain (SIA, 2016). The industry is mature, with most segments dominated by a small number of firms located in the US, South Korea, Taiwan, Japan, Europe, and China. There are considerable entry barriers, most notably first mover technology advantages, intellectual property, and extremely high fixed costs (King, 2003). Competitive advantage in the semiconductor industry is dependent on the manufacturing process, which greatly influences performance, power consumption, time to market, and cost. Due to its complex nature, profitable semiconductor manufacturing requires large-scale operations and an imperative to fully utilize capacity.

To facilitate the commercialization of new digital technologies and the utilization of available capacity, the industry disintegrated into the specialist segments of design, fabrication, assembly, testing, and packaging, as described above. These distinct activities form a global value chain where both down- and upstream firms can generate value through innovation. The suppliers of materials, EDA software, IP blocks, and manufacturing equipment complement the core firms in the value chain to form the geographically distributed semiconductor ecosystem. There are still two parallel operational strategies in the semiconductor industry. The traditional mode of operation is being an IDM that vertically integrates design, manufacturing, test, and assembly. Within the newer fabless-foundry model that emerged with industry fragmentation, specialized firms cooperate in the ecosystem. The semiconductor value chain is presented in Figure 2 below.

Modularity in both product architecture and industrial organization provides strategic opportunities for entrants and incumbents in the electronics industry (Sturgeon & Kawakami, 2010). Additionally, offshoring to exploit lower labor costs and gain better access to growing Asian markets has contributed to semiconductor and electronics manufacturing shifting to Asia. To summarize, the semiconductor industry is characterized by rapid technological advances, global markets, and strategically designed industrial policies (Flamm, 2010).

Figure 2 The semiconductor value chain



An adaptation from SIA (2016).

3.4 Industry challenges to meeting diverse computational demands

Taking a top-down perspective, the exponential growth of data and emerging technologies—such as autonomous vehicles, 5G, the internet of things, and scientific computing—promote the demand for faster and more energy-efficient computers. Ranging from data centers to IoT edge devices, each technology has different requirements for the underlying semiconductor ICs. As an example, we can contrast the systems on a chip (SoC) used in smartphones that integrate the central processing unit (CPU), graphics processing unit (GPU)—as well as network, video, and AI processing—on a single silicon die with the large monolithic CPU designs used in data centers (Waldrop, 2016).

There are myriad technical details about advancing semiconductor manufacturing, and progress is needed in materials, transistor design, manufacturing, packaging technologies, and microarchitectures. Extensive coordination between designers, materials suppliers, equipment makers, and manufacturers is needed in order to realize these goals (Waldrop, 2016). We briefly dive into the lowest layers of the stack to give an outlook on how the semiconductor industry plans to meet the insatiable demand for more computation.

At the heart of the microprocessors and memory devices in our computers is the transistor, billions of which have been integrated in modern ICs. Improving the performance and boosting the density of transistors has been the most straightforward way to speed up and cheapen all the digital devices we use today. Although the shape and materials of transistors have changed, the same basic structure of complementary metal oxide semiconductor (CMOS) technology (a gate controlling an electric channel between the transistor's source and drain), which was invented in the 1960s,[7] remains in place today. This is the premise of the empirical observation made by G. Moore in 1965 and has been sustained by the semiconductor industry for over 50 years (Ye, Ernst, & Khare, 2019). But the scaling of CMOS transistors has continually faced physical challenges and will eventually come to an end (Waldrop, 2016).

From a bottom-up perspective (e.g., considering what type of transistor is used), the industry has made multiple transitions throughout history (O'Reagan & Fleming, 2018). The most recent and relevant shift was the adoption of the fin field-effect transistor (FinFET), a new transistor type which was technically proven around 2001and first commercialized in 2011 by Intel. The transition required a concerted collaboration between major American semiconductor companies, leading-edge universities, and federally funded research programs. Interestingly, the FinFET was successful because it was not too radical a change. Because of immense investments by the international semiconductor industry in CMOS technologies, the FinFET needed to fit within the existing manufacturing paradigm (O'Reagan & Fleming, 2018).

The FinFET breakthrough has sustained Moore's law during the 2010s; however, TSMC and Samsung have announced that they will transition to nanosheet transistors at the leading edge in two to three years. FinFETs suffer from electrical leakage that becomes untenable at the upcoming 3 nm[8] node. Nanosheet transistors wrap the gate around the channel to provide better electrodynamic control over the transistor channel, a concept that researchers have tried to utilize since as early as 1990[9] (Ye, Ernst, & Khare., 2019). This highlights the long development cycle in the bottom layers of the technology stack: 30 years from conceptual idea to the start of mass production.

Extreme ultraviolet (EUV) lithography equipment, introduced at the 7 nm node and solely provided by Dutch company ASML,[10] has allowed single exposure patterning of critical chip structures in the advanced nodes. Single patterning provides cost, yield, and cycle-time benefits in manufacturing. But beyond the 5 nm node, multi-patterning EUV lithography becomes inevitable, which adds to the wafer costs. Lithography equipment therefore needs improving in order to shift back to the single exposure patterning of critical chip features at future (1 nm) nodes (Samavedam, 2020).

Further problems are caused by the fact that the amount of heat a microprocessor can dissipate (i.e., the power density) has not scaled in the past decade. Processor clock rates are being kept down to manage heat and the industry has thus shifted to multicore microarchitectures to utilize increasing transistor counts. Many workloads can be parallelized to take advantage of many processor cores and reach a solution as quickly as a faster single processor core. One solution for the heat issue is to introduce new materials in the channel region of the transistor, which has the potential to reduce heat and provide higher efficiency (Ye, Ernst, & Khare, 2019).[11] With nanosheet transistors, improvements in manufacturing equipment and new materials, transistor density can continue to scale for eight to ten years but performance increase at fixed power will be likely to slow down (Samavedam, 2020).

Despite the potential to increase performance through various innovations, the increasing complexity of sustaining Moore's law has led to rising costs in both fabrication and design. A leading-edge fabrication plant now costs over $15 billion (TSMC, 2018) and non-recurring engineering work on a 7 nm microarchitecture (the currently maturing manufacturing process) reportedly costs up to $300 million (Lapedus, 2018). The huge costs of regenerating manufacturing infrastructure for the technology successors will most likely constrain the future of the industry (Isaac, 1997). The industry is hence shifting towards heterogeneous integration with die-to-die connectivity as a cost-efficient way to improve system performance (Samavedam, 2020).

3.5 The US leads semiconductor value capture and creation

The US is a clear leader both in creating and capturing value in the semiconductor industry. We use sales revenue and R&D expenditure figures in support of this claim.

Global semiconductor sales were $481 billion in 2018 and annual sales growth is forecast at 4.6% through 2022.[12] The growth in demand is driven by high-performance computing, electric and autonomous vehicles, and the proliferation of AI applications, as well as by the implementation of 5G networks around the globe (PWC, 2019). On the other hand, declining PC and laptop sales, as well as stagnated smartphone sales, create a drag on semiconductor demand. Investment is generally driven by demand for technologically superior products with improved capabilities and reliability (SIA, 2016).

US headquartered firms account for 47% of revenue while firms headquartered in China only generated 7% of global revenues. Revenue generated in the semiconductor industry by region and across segments in 2019 is presented in Table 2 below. Fabless design firms and IDMs are included in the same category since they both have chip design capabilities. The IDM and design segment is by far the largest in semiconductor value chains and includes multiple companies with revenues exceeding $20 billion (e.g., the IDMs Intel, Samsung, SK Hynix, and Micron, as well as the fabless companies Qualcomm and Broadcom). The US has a particularly strong position in chip design, IDMs, manufacturing equipment, and EDA software. China, on the other hand, has a relatively large share of outsourced assembly and testing but still lags far behind the US and other advanced semiconductor countries in other segments.

Table 2 Semiconductor industry sales by region and segment[13]

Semiconductor industry sales, in billions				
	Total	US	China	The rest of the world
IDM & design	$407.7	54%	7%	39%
Equipment	$71.6	47%	2%	52%
Foundry	$54.7	11%	8%	81%
OSAT	$28.3	14%	21%	64%
IP & EDA	$9.5	78¤	1%	21%

van Hezewijk (2020).

To indicate the relative positions of the countries participating in semiconductor value chains, we present consolidated data on industry and government R&D expenditure in Table 3 on next page. R&D expenditure in the semiconductor industry has averaged 15% of sales (SIA, 2016) and we see it as a proxy for value creation. US-based semiconductor companies account for over half of this investment. China is the outlier with the government providing most of the R&D funding.[14] However, a large share of Chinese government investment is allocated to capacity installments and acquiring existing technology (van Hezewijk, 2019), which only upgrade local capabilities incrementally. Finally, honorable mentions go to South Korea and Taiwan, as well as to the Netherlands, whose research and investments have made the continuation of Moore's law possible.[15]

Table 3 Geographical distribution of semiconductor R&D expenditure

Semiconductor R&D expenditure, in billions						
	US	China	The rest of the world			
			S. Korea	Taiwan	Japan	Netherlands
Revenue	$270.9	$41.3	$80.9	$75.9	$50	$25.4
Industry R&D	$37.8	$2.6	$8.4	$6.8	$5.2	$3.6
Government R&D	$1.5	$5.5	$1.7	n/a	n/a	$0.1
% of revenue	15%	20%	12%	9%	10%	15%
% of total	54%	11%	14%	9%	7%	5%

van Hezewijk (2019).

4 An overview of Sino-American semiconductor policy

4.1 The USA—maintaining leadership

As seen above, the US holds dominant market positions in the EDA, equipment, and IDM/design segments of the semiconductor industry. But China is the largest IC market globally and US semiconductor firms generate 36% of their revenue in the mainland market (Fitch & Davis, 2020). The largest equipment firms and IDMs generate over twice as much revenue in China, as opposed to the US, highlighting the importance of the Chinese market (van Hezewijk, 2019). Any (US or Chinese) policy that diminishes American companies' revenue from China will hurt their competitiveness.

Adding to American woes, Intel, which was once the paragon of advanced chipmaking is now one process generation between behind TSMC and has announced delays in developing its most advanced 7 nm[16] manufacturing process (Salter, 2020). The US thus finds itself amid a technology war with China at a point when its domestic semiconductor mass manufacturing capabilities are beginning to trail behind the leading edge. Nevertheless, US industry, academia, and the US government are again collaborating to tackle cost, complexity, and competitive challenges with a similar model to the FinFET breakthrough discussed above (DARPA, 2020).

US lawmakers have realized that the domestic semiconductor industry's competitiveness and investment capacity may be diminished by the trade war and have proposed legislation that would provide over $20 billion in aid to support US semiconductor manufacturing (Nellis, 2020). The bill would provide investment tax credits, a federal "matching" fund to match state incentives, allocate federal funds for semiconductor R&D, and also focus on developing advanced IC packaging capabilities (Warner, 2020).

TSMC has been enticed to build a $12 billion semiconductor foundry in Arizona and has reportedly agreed on subsidies with the local government in order to offset higher production costs[17] in the US (Wu, 2020). However, the planned foundry capacity is small compared with TSMC's Taiwanese "giga fabs" and the manufacturing process would be one generation old upon completion. But TSMC's Arizona fab could be trusted for US defense applications with smaller production runs.

As witnessed in the recent trade war escalation, the US evidently has the technological clout to inflict damage on Chinese firms and thus restrict China's technological development. It is not the first time the US has restricted semiconductor exports to China. For instance, Intel was denied an export license to supply Xeon server-grade processors to a Chinese supercomputer in 2015, citing concerns over nuclear device development (BBC, 2015). Another example of protectionist measures by the US Commerce Department was the banning of all exports of components and software to the second-largest Chinese telecom equipment firm, ZTE, in 2018. Restrictions were imposed because ZTE failed to comply with economic sanctions against Iran and North Korea. A settlement requiring ZTE to pay a $1 billion fine was reached and the ban was subsequently removed. However, ZTE is said to remain under close scrutiny by US authorities (Ballentine, 2018).

US prosecutors have furthermore indicted Taiwanese foundry UMC, as well as newly established Chinese memory producer Fujian Jinhua, of stealing the trade secrets of Micron, a US DRAM manufacturer. A manager became part of Micron following an acquisition and then became a president of Micron's Taiwan subsidiary MMT. The manager resigned from MMT after two years, bringing with him some 900 proprietary files when he joined UMC in 2015. A partnership was then quickly established with Fujian Jinhua to transfer DRAM technology for mass production. Other engineers from MMT brought more intellectual property with them when they were recruited to UMC (Department of Justice, 2018).

4.2 China—catching up and securing access

China is extremely dependent on semiconductor imports. The import value was $312 billion in 2018, amounting to over 60% of global sales (*The Economist*, 2020c). Recent events in the trade war underscore China's predicament—it is subject to politically motivated decisions across the Pacific, and Huawei finds itself effectively cut off from the leading-edge chip supply. China is playing technological catch up in the semiconductor industry while it is trying to secure its supply.

China has launched *Guidelines to Promote a National Integrated Circuit Industry* in 2014 and Made in China 2025 in its latest efforts to achieve technological self-sufficiency (VerWey, 2019). The country has implemented industrial policies since the 1960s to support the strategic development of a domestic semiconductor industry.

Made in China 2025 outlines a vision to "develop the IC design industry, speed up the development of the IC manufacturing industry, upgrade the advanced packaging and testing industries, facilitate breakthrough in the key equipment and materials of integrated circuits." Furthermore, Made in China 2025 describes aiming to domestically produce 70% of chips by 2025. The above-mentioned guidelines called for $150 billion to be invested by 2025 and set out a two-pronged strategy that focuses on outbound investments in foreign technology companies and the facilitation of domestic greenfield investment and joint ventures (VerWey, 2019).

Between 2014 and 2017, Chinese investments in US semiconductor companies totaled a record $10 billion (Yue & Lu, 2017). But Chinese acquisitions of foreign technology firms have now become subject to scrutiny. The acquisition of US firm Lattice Semiconductor was blocked on national security grounds (Executive Order, Sep 13th, 2017). Furthermore, Germany has introduced new measures that allow the government to scrutinize and block deals in strategic economic areas (e.g., in AI, robotics, semiconductors, biotechnology, and quantum technology) (Chazan, 2019).

The current well-funded and clearly defined policy is part of a continued effort by the Chinese government to promote nationalism and achieve independence from foreign technology (Zenglein & Holzmann, 2020). The government has enacted the strategy by establishing the China Integrated Circuit Industry Fund, which raised $22 billion in 2014 and $29 billion in 2019 (van Hezewijk, 2019). This centrally established "big fund" guides provincial governments in their efforts to implement the industrial policy, and a United States Trade Representative section 301 report (2018, p. 94) cites an SIA estimate that provincial and municipal IC funds have raised an additional $80 billion since 2014.

China is championing SMIC to pursue the goals set out in Made in China 2025. The Shanghai-based foundry raised close to $10 billion in financing in the spring of 2020 in order to increase capacity and develop its manufacturing processes (Wei, 2020; Fang & Li, 2020). Higher up in the value chain, Chinese chip design companies, such as HiSilicon and Tsinghua Unigroup, are among the global top 10 IC design firms by revenue (*The Economist*, 2018). China continually tries to recruit engineers from Taiwan by offering better compensation (Ihara, 2019; Fang, 2020).

However, the country is set to fall far short of the targets set out in Made in China 2025, calling into question the efficiency of the centrally designed incentives in the Chinese approach. Hybrid firms (Chinese enterprises with foreign financing) have furthermore been the most innovative in developing Chinese technology when compared with local state-owned enterprises (Fuller, 2016). Looking at Chinese IC production, domestic fabricators (those with HQ in China) covered only 5% of DRAM, flash memory, and logic sales in China in 2018. Accounting for both domestic and foreign producers, ICs fabricated in China covered 15% of the demand (IC insights, 2019).

New tax subsidies for semiconductor companies were announced in August of 2020 (Kharpal, 2020). Chinese efforts have so far merely had incremental success because of the industry's highly globalized, competitive, and market-driven nature. Companies absolutely need more than cash to compete and Chinese policy looks likely to only have a marginal impact on Chinese semiconductor firm's ability move up value chains.

China leverages the size of its domestic market in its soft power retaliation. For instance, Qualcomm's merger with NXP fell through in 2018 as China's State Administration for Market Regulation (SAMR) was delaying approval of the deal (Martina & Nellis, 2018). Two thirds of Qualcomm's revenue are generated in China, and it thus needed Chinese approval of the acquisition. In the same year, SAMR started an investigation against Samsung, SK Hynix, and Micron for price-fixing in DRAM markets. The three firms collectively control a daunting 95% market share (Harris, Jung-a, & Song, 2018). China again used access to its domestic markets as leverage, but it is not the first to punish DRAM producers for price-fixing. Both Samsung and SK Hynix have paid hundreds of millions in fines for price-fixing to both the European Commission and the US Justice department in 2010 and 2005 respectively.

5 Discussion and policy implications

5.1 The next semiconductor crisis and technological separation

Brown and Linden (2011) argued that different interconnected and recurring "crises" shape the semiconductor industry. Sturgeon (2011) saw the economic crisis of 2008–2009 as an inflection point at which Asian firms assumed a leading role in developing the global electronics industry. Building on these commentaries, we see that the ongoing Sino-American technology separation marks an inflection point for global competition in the semiconductor industry—it forces change in value chains and innovation networks.

To recapitulate, the American semiconductor industry is faced with a crisis of increased competition and the loss of leading-edge manufacturing. On the other hand, China's semiconductor industry faces a limited supply of experienced engineers and risks being cut off from critical American and European manufacturing equipment by decades-long technology barriers. Moreover, Chinese technology products face a branding crisis in Western markets, and they are seen as being unsecure and under the malign influence of the Communist party. The crisis is compounded by a pandemic-induced recession.

The Sino-American technology separation might result in two separate industrial ecologies and two technological spheres of influence. Defensive American action will slow Huawei's progress. On the other hand, actions taken to limit the supply of leading-edge chips absolutely reinforces China's drive to technology self-sufficiency. We (i.e., all blocs) should also harbor no illusions about the costs of a technology separation at the lowest levels of the hardware stack. Many American and European semiconductor firms have their largest markets in China and might be greatly affected by further escalations in the conflict. In response to Chinese state-led upgrading, the US is drafting a bill that would provide tens of billions of dollars in support to the US semiconductor industry.

5.2 Technical challenges and national policies

Technical challenges to meeting increased demand for computational power affects the top layers of the technology stack as well. With the increasing cost and complexity to sustain Moore's law, semiconductor research institutes now explore other computational technologies—such as quantum, neuromorphic,[18] or photonic computing—that might provide solutions in the medium term (Lapedus, 2019). The ultimate question is how a balance can be struck between investments in current and future needs. Investing in mathematics, algorithms, and computer science is as important as developing new types of logic devices and manufacturing techniques (PRACE, 2018). Although the industry is vertically specialized, we observe that platform and system companies expand vertically into chip design for strategic reasons.[19]

The diffusion of technical semiconductor capabilities and expected changes in technology have led to the establishment of state-sponsored national champions that directly engage in fierce global competition, resulting in high-stakes political intervention (Flamm, 2010). It is simply not possible to completely stop the diffusion of technology, and protecting the leadership status of a strategically important industry such as semiconductors requires deep collaboration, a focus on IP protection, bringing new innovations to market, and setting standards. Competing in global semiconductor markets is not cheap or easy because products are founded on long scientific research projects and some segments of development are protected by national security priorities (PCAST, 2017). Additionally, industrial policy has frequently supported the establishment of local semiconductor businesses (PCAST, 2017). This has implications for trade and industrial policy, which needs to account for the reality of supranational blocs investing in new technology that disrupts industrial landscapes.

5.3 Policy implications for Europe

Any public policy aiming for technology sovereignty should consider the limited talent pool, market development, innovative capabilities, national research priorities, and new competition (Ernst, 2010). Europe clearly needs deep external collaboration in order to keep abreast with semiconductor innovation abroad. Simultaneously looking inward to improve European cooperation is likely to be required in order to succeed.

Given the dichotomy of a technology separation, the options for Europe can be outlined as follows:

- **Choose American technology:** Continue participating in leading American innovation networks and be a fast adopter of US technology products in order to quickly reap the benefits of high-risk, low-return US investments. The main question with this strategy is if American interests curb European decision-making autonomy.
- **Choose Chinese technology:** Chinese hardware is not extensively used in Europe, but systems and platforms are, in principle, available to Europeans. Adopting Chinese technology might become necessary in order to access the main growth market for MNCs, but is all business good business?
- **Upgrade European technology:** In theory Europe has an option for ambitious industrial upgrading in semiconductor manufacturing with globally recognized research institutes Cea-Leti and Imec, and dominant lithography supplier ASML, as well as the IDMs NXP, STMicroelectronics, and Infineon. This strategy, however, requires a commitment of 20–30 years, as well as multi-billion-euro funding programs. A technology leadership strategy is extremely costly, and a more prudent option might be to diversify into multiple technology areas (see Ernst, 2010).

5.4 Concluding remarks

It is currently unclear what Europe's strategy is in regard to reacting to the changing competitive landscapes in the semiconductor industry. If a commitment to any option above is to be made on a European level, Finland's policy of technology neutrality and standards might become obsolete quickly. From the perspective of Finland's export-dependent economy, the risks of losing global sales opportunities needs to be considered when planning for the industrial and digital future of Europe. In the future businesses might be forced to become more flexible in terms of their product designs, for example that Chinese hardware and software must be used in products for the Chinese market. If the world is moving towards unilateralism, Europe should definitively consider how to keep in contact with regional innovation networks in Silicon Valley, Japan, South Korea, and Taiwan, but also China.

Today the semiconductor industry is facing a crisis that is likely to accelerate innovative efforts within supranational blocs. Current broad disagreements in international relations, alarming as they are, heighten the risk of uneven development in different parts of the world. Rapidly evolving technology and digitization will continue driving large shifts in the social and economic order. Therefore, it is hard to tell if there will be a winning side or standard in the current technology confrontation or if new rules for international technology competition and collaboration will emerge to accommodate multiple actors.

Endnotes

1 Globalization is commonly used to frame discussions in social sciences, politics, business management, and journalism.

2 Mainly Taiwan, South Korea, & Japan.

3 The TSMC is the technology leader in advanced semiconductor manufacturing and commands a majority share of the IC foundry market. Huawei and Apple generated 14% and 23% of the TSMC's revenues in 2019 respectively, highlighting American and Chinese dependence on TSMC's leading-edge manufacturing in order to deliver products with superior performance (TSMC, 2020). China is championing Semiconductor Manufacturing International Corporation, which is 1/10th the size of the TSMC, to spearhead efforts of semiconductor self-sufficiency. The Shanghai-based foundry raised close to $10 billion in financing in the spring of 2020 in order to increase capacity and develop its manufacturing processes (Wei, 2020; Fang & Li, 2020). Meanwhile leading-edge semiconductor manufacturing in the US is facing headwinds as Intel has announced delays in its upcoming 7-nm process node (Alcorn, 2020).

4 *Sovereignty* in this context relates to either common European values, maintaining control over the technology used in member states, building competitiveness of European MNCs, or improving cybersecurity.

5 This raises multiple other research questions, e.g., With Sino-American relations souring, can Europe remain neutral regarding the digital technology stack? Will Europe be forced to choose between American or Chinese technology? Should Europe invest more resources in developing hardware? What is the European position on semiconductors? And what are the European policy responses?

6 Concrete examples include €80 million in EU seed funding for the European Processor Initiative, an industry consortium that will design a high-performance computing processor (Cordis Europa, 2020; EPI, 2020), and €3 billion in EU investments in high-performance computing resources (EuroHPC, 2020).

7 State-funded academic research has been vital since the formative days of the semiconductor industry and defense spending was a catalyst for early growth (O'Reagan & Fleming, 2018).

8 The naming convention is a heritage from planar transistors; however, the current node names do not have a direct relation to the size of physical features but rather only reflect the degree of transistor miniaturization.

9 Additionally, nanosheets provide flexibility since the width the sheet can be varied to either boost performance or limit power consumption.

10 See ASML's equipment in the work of Seeker (2019).

11 Making transistors, e.g., from III–V semiconductors with higher electron mobility.

12 Semiconductor sales statistics should be compared with care as they risk double counting.

13 Consolidated data contains more (and smaller) companies from China but only the largest and most important companies for the US and the rest of the world. The materials segment is excluded.

14 Chinese government spending is calculated from the first tranche of the national IC fund, spread out over a five-year investment period.

15 South Korea, Taiwan, Japan, and the Netherlands together make up 90% of revenue generated in the rest of the world and are therefore included in the comparison on R&D expenditure.

16 Intel's 7 nm process is like TSMC's 5 nm process in terms of transistor density.

17 Technician salaries are 2 times higher in the US compared with Taiwan, and a lack of assembly, test, and other ancillary services raises costs (Patterson, 2020).

18 In practice, neuromorphic computing has meant AI accelerators that parallelize the training task in hardware.

19 In the US, this means a concentration of semiconductor talent in "tech giants." The US hardware industry might face a shortage of skilled labor due to software engineering work having stronger "pull."

References

Alcorn, P., 2020. "Intel's 7nm is Broken, Company Announces Delay Until 2022, 2023" [online]. Tom's Hardware, Jul 18th. Available at: https://www.tomshardware.com/news/intel-announces-delay-to-7nm-processors-now-one-year-behind-expectations (accessed 29 July 2020)

Amsden A.H., 1989. "Asia's next giant: South Korea and late industrialization". Oxford University Press on Demand.

Assche A.v. & Gangnes, B., 2010. "Electronics production upgrading: is China exceptional?". Applied Economics Letters, 17:5. p. 477-482.

Baldwin, R.E., 2006. "Globalization: the great unbundling(s)". Economic Council of Finland.

Ballentine, C., (2018). "U.S. Lifts Ban That Kept ZTE From Doing Business With American Suppliers" [online]. The New York Times, Jul 13th. Available at: https://www.nytimes.com/2018/07/13/business/zte-ban-trump.html (accessed 20 July 2020)

BBC, 2015. "US nuclear fears block Intel China supercomputer update" [online]. BBC, April 10th. Available at: https://www.bbc.com/news/technology-32247532 (accessed 20 July 2020)

Brown, C. & Linden, G., 2005. "Offshoring in the Semiconductor Industry: A Historical Perspective". In *Brookings Trade Forum*, p. 279-333. Brookings Institution Press.

Brown, C. & Linden, G., 2011. "Chips and change". MIT Press.

Chazan, G., 2019. "Germany acts to stop sale of tech companies to non-EU investors" [online]. Financial Times, Nov 29th. Available at: https://www.ft.com/content/d0964cd0-12af-11ea-a7e6-62bf4f9e548a (accessed 3 August 2020)

Cherney M.A., 2020. "How Broadcom May Have Helped Make Google's Latest AI Chip". Barron's. Aug 12th. Available online at: https://www.barrons.com/articles/how-broadcom-may-have-helped-make-googles-latest-ai-chip-51597267076 (accessed 13 August 2020)

Cordis Europa, 2020. "On the road to exascale computing". EU research results. Available at: https://cordis.europa.eu/project/id/826647 (accessed 3 July 2020)

DARPA (Defense Advanced Research Projects Agency), 2020. "DARPA Electronics Resurgence Initiative" [online]. Website. Available at: https://www.darpa.mil/work-with-us/electronics-resurgence-initiative (accessed 8 August 2020)

Department of Commerce, 2019. "Department of Commerce Announces the Addition of Huawei Technologies Co. Ltd. to the Entity List" [online]. Press release, May 15th. Available at: https://www.commerce.gov/news/press-releases/2019/05/department-commerce-announces-addition-huawei-technologies-co-ltd (accessed 3 June 2020)

Department of Commerce, 2020. "Commerce Addresses Huawei's Efforts to Undermine Entity List, Restricts Products Designed and Produced with U.S. Technologies" [online]. Press release, May 15th. Available at: https://www.commerce.gov/news/press-releases/2020/05/commerce-addresses-huaweis-efforts-undermine-entity-list-restricts (accessed 3 July 2020)

Department of Justice, 2018. "PRC State-Owned Company, Taiwan Company, and Three Individuals Charged With Economic Espionage" [online]. Department of Justice release, Nov 1st. Available at: https://www.justice.gov/opa/pr/prc-state-owned-company-taiwan-company-and-three-individuals-charged-economic-espionage (accessed 10 July 2020)

Dick, A.R., 1996. "Comment" [Peer commentary on the paper "Trade policies and the semiconductor industry" by D.A. Irwin]. The political economy of American trade policy, p. 11-72. University of Chicago Press.

The Economist, 2018. "The semiconductor industry and the power of globalization" [online]. The Economist, Dec 1st. Available (behind paywall) at: https://www.economist.com/briefing/2018/12/01/the-semiconductor-industry-and-the-power-of-globalisation (accessed 10 June 2020)

The Economist, 2020a. "How big is China's economy? Let the Big Mac decide" [online]. The Economist, July 15th. Available (behind paywall) at: https://www.economist.com/graphic-detail/2020/07/15/how-big-is-chinas-economy-let-the-big-mac-decide (accessed 20 July 2020)

The Economist, 2020b. "TikTok and the Sino-American tech split". The Economist. July 11th. Available (behind paywall) at: https://www.economist.com/leaders/2020/07/11/tiktok-and-the-sino-american-tech-split (accessed 20 July 2020)

The Economist, 2020c. "China is slowly moving up the microprocessing value chain" [online]. The Economist, Jan 2nd. Available (behind paywall) at: https://www.economist.com/technology-quarterly/2020/01/02/china-is-slowly-moving-up-the-microprocessing-value-chain (accessed 10 June 2020)

EPI, 2020. "European Processor Initiative". Website. Available at: https://www.european-processor-initiative.eu/project/epi/ (accessed 3 July 2020)

Ernst, D., 2010. "Upgrading through innovation in a small network economy: insights form Taiwan's IT industry". Economics of Innovation and New Technology, 19(4), p. 295-324.

EuroHPC, 2020. "EuroHPC: Leading the way in European Supercomputing". Website. Available at: https://eurohpc-ju.europa.eu/ (accessed 3 July 2020)

European Commission, 2020. "WHITE PAPER: On Artificial Intelligence – A European approach to excellence and trust" [online]. 19th Feb. Available at: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (accessed 3 July 2020)

Fang, C.T., 2020. "China hires over 100 TSMC engineers in push for chip leadership" [online]. Nikkei Asian Review, Aug 12th. Available at: https://webcache.googleusercontent.com/search?q=cache:NsKyrk4lXzIJ:https://asia.nikkei.com/Business/China-tech/China-hires-over-100-TSMC-engineers-in-push-for-chip-leadership+&cd=2&hl=en&ct=clnk&gl=fi&client=firefox-b-d (accessed 13 August 2020)

Fang, C.T. & Li, L., 2020. “Chipmaker SMIC touts benefits of $6.5 billion China share listing plan” [online]. Nikkei Asian Review, Jul 5th. Available at: https://webcache.googleusercontent.com/search?q=cache:fnSw6Z9uFHMJ:https://asia.nikkei.com/Business/China-tech/Chipmaker-SMIC-touts-benefits-of-6.5bn-China-share-listing-plan+&cd=1&hl=en&ct=clnk&gl=fi&client=firefox-b-d

Fitch, A. & Davis, B., 2020. “U.S. Chip Industry Fears Long-Term Damage From China Trade Fight” [online]. Wall Street Journal, Mar 9th. Available at: https://www.marketscreener.com/ON-SEMICONDUCTOR-CORPORAT-10340/news/U-S-Chip-Industry-Fears-Long-Term-Damage-From-China-Trade-Fight-30129271/

Flamm, K., 2010. “Mismanaged Trade? Strategic Policy and the Semiconductor Industry”. Brookings Institution Press.

Frankel, J., 2020. “Is China overtaking the US as a financial and economic power?”. The Guardian. May 29th. Available at: https://www.theguardian.com/business/2020/may/29/is-china-overtaking-the-us-as-a-financial-and-economic-power (accessed 20 July 2020)

Fransman, M., 1990. “US-Japan Science and Technology Exchange: Patterns of Interdependence”. Westview Press.

Fuller, D.B., 2016. “Paper Tigers, Hidden Dragons”. Oxford University Press.

Harris, B., Jung-a, S. & Feng, E., 2018. “China alleges ‘massive’ evidence of chipmaker violations” [online]. Financial Times, Nov 19th. Available at: https://www.ft.com/content/d626833a-ebb5-11e8-89c8-d36339d835c0 (accessed 10 July 2020)

Harrison, A. 2014. “Trade and industrial policy”. The Oxford Companion to the Economics of China, p. 161.

Hirst, P., Thompson, G. & Bromley, S., (2015). “Globalization in Question”. John Wiley & Sons.

IC insights, 2019. “Can We Believe The Hype About China’s Domestic IC Production Plans?”. Excerpt from market report. Available at: https://semiwiki.com/forum/index.php?threads/can-we-believe-the-hype-about-china%E2%80%99s-domestic-ic-production-plans.11398/ (accessed 3 July 2020)

Ihara, K., 2019. “Taiwan loses 3000 chip engineers to ‘Made in China 2025’” [online]. Nikkei Asian Review, Dec 3rd. Available (behind paywall) at: https://asia.nikkei.com/Business/China-tech/Taiwan-loses-3-000-chip-engineers-to-Made-in-China-2025 (accessed 3 July 2020)

Irwin, D.A., 1996. “Trade policies and the semiconductor industry”. The political economy of American trade policy, p. 11-72. University of Chicago Press.

Isaac, R., 1997. “Viewpoint: beyond silicon ... and back again”. IEEE Spectrum, 34(1).

Kenney, M. & Pon, B., (2012). “Structuring the smartphone industry: is the mobile internet OS platform the key?”. Journal of industry, competition, and trade, 11(3), p. 239-261.

Kenney, M., Breznitz, D., & Murphree, M., 2013. “Coming back home after the sun rises: Returnee entrepreneurs and growth of high tech industries”. Research Policy, 42(2), p. 391-407.

Kharpal, A., 2019. “Alibaba unveils its first A.I. chip as China pushes for its own semiconductor technology”. CNBC. Sep 19th. Available at: https://www.cnbc.com/2019/09/25/alibaba-unveils-its-first-ai-chip-called-the-hanguang-800.html (accessed 11 August 2020)

Kharpal, A., 2020. “China unveils policies to boost chipmakers as tensions with U.S. rise. Analysts say they may not help” [online]. CNBC, Aug 10th. Available at: https://www.cnbc.com/2020/08/11/china-policies-to-boost-chipmakers-as-tensions-with-us-rise.html (accessed 11 August 2020)

King, D., 2003. “Size Matters: Barriers to Entry in the Microelectronics Industry”. National Defense University.

Koskinen, M., 2020. “Kiinan ja Yhdysvaltain siteet purkautuvat – mihin kaikkeen suuri ero voi johtaa?”. Ulkopolitiikka. May 22nd. Available in Finnish at: https://ulkopolitiikka.fi/lehti/2-2020/kiinan-ja-yhdysvaltain-siteet-purkautuvat-mihin-kaikkeen-suuri-ero-voi-johtaa/ (accessed 5 August 2020)

Lapedus, M., 2018. ”Big Trouble At 3nm” [online]. Semiconductor Engineering, June 21st. Available at: https://semiengineering.com/big-trouble-at-3nm/ (accessed 3 August 2020)

Lapedus, M., 2019. ”What’s Next For AI, Quantum Chips” [online] Semiconductor Engineering, Jan 24th. Available at: https://semiengineering.com/whats-next-for-ai-quantum-chips/ (accessed 20 July 2020)

Larsen, M.M., Seppälä, T. & Ali-Yrkkö, J., (2018). “The changing geography and ownership of value creation: evidence from mobile telecommunications”. Industry and innovation, 25(7), p. 675-698.

Martina, M. & Nellis, S., 2018. “Qualcomm ends $44 billion NXP bid after failing to win China approval” [online]. Reuters, July 25th. Available at: https://www.reuters.com/article/us-nxp-semicondtrs-m-a-qualcomm/qualcomm-ends-44-billion-nxp-bid-after-failing-to-win-china-approval-idUSKBN1KF193 (accessed 10 July 2020)

Mathews, J.A., 1997. “A Silicon Valley of the East: Creating Taiwan’s Semiconductor Industry”. California Management Review, 39(4), p. 26-54.

Nellis, S., 2020. “U.S. lawmakers propose $22.8 billion in aid to semiconductor industry” [online]. Reuters, Jun 10th. Available at: https://www.reuters.com/article/us-usa-semiconductors/u-s-lawmakers-propose-22-8-billion-in-aid-to-semiconductor-industry-idUSKBN23H39M (accessed 10 July 2020)

Nenni, D., 2020. “Dr. Morris Chang Wiki” [online]. SemiWiki.com. Jan 1st. Available at: https://semiwiki.com/wikis/industry-wikis/dr-morris-change-wiki/ (accessed 3 August 2020)

Nolan, P. 2001. “China and the global economy: national champions, industrial policy and the big business revolution”. Springer.

Patterson, A., 2020. “TSMC Reiterates Support for Arizona Fab” [online]. EETimes Asia, Jun 22nd. Available at: https://www.eetasia.com/tsmc-reiterates-support-for-arizona-fab/ (accessed 10 July 2020)

PCAST (President's Council of Advisors on Science and Technology), 2017. "Ensuring Long-Term U.S. Leadership in Semiconductors" (online). Available at: https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_ensuring_long-term_us_leadership_in_semiconductors.pdf (accessed 3 August 2020)

PRACE (Partnership for Advanced Computing in Europe), 2018. "The Scientific Case for Computing in Europe 2018-2026". Insight Publishers. Available at: https://prace-ri.eu/third-scientific-case/ (accessed 10 July 2020)

Prestowitz, C.V., 1988. "Trading places: How we allowed Japan to take the lead".

PWC (PricewaterhouseCoopers), 2019. "Opportunities for the global semiconductor market: Growing market share by embracing AI". Consultancy report, April 3rd. Available at: https://www.pwc.com/gx/en/industries/tmt/publications/global-tmt-semiconductor-report-2019.html (accessed 10 July 2020)

Rostami, M., Koushanfar, F., & Karri, R., 2014. "A primer on hardware security: Models, methods, and metrics". Proceedings of the IEEE, 102(8), 1283-1295.

Salter, J., 2020. "As 7nm schedule continues slipping, Intel contemplates 3rd-party fabs" [online]. arstechnica, July 24th. Available at: https://arstechnica.com/gadgets/2020/07/as-7nm-schedule-continues-slipping-intel-contemplates-3rd-party-fabs/ (accessed 3 August 2020)

Samavedam, S., 2020. "Five trends that will shape the future semiconductor technology landscape" [online]. Semiconductor Digest. July. Available at: https://www.imec-int.com/en/articles/five-trends-will-shape-future-semiconductor-technology-landscape (accessed 3 August 2020)

Saxenian, A.L., 2006. "The New Argonauts: Regional Advantage in a Global Economy". Harvard University Press.

SIA (Semiconductor Industry Association), 2016. "Beyond Borders: The global Semiconductor Value Chain".

Seeker, 2019. "The Extreme Physics Pushing Moore's Law to the Next Level" [online]. Video. Available at: https://www.youtube.com/watch?v=f0gMdGrVteI (accessed 3 August 2020)

Seppälä, T., Kenney, M., & Ali-Yrkkö, J., 2014. "Global supply chains and transfer pricing: Insights from a case study". Supply Chain Management, 19(4), p. 445-454.

Sturgeon, T., & Kawakami, M. (2010). "Global value chains in the electronics industry: was the crisis a window of opportunity for developing countries?". Global Value Chains in a Postcrisis World, Washington, DC: The World Bank. p. 245-301.

TSMC, 2018. "TSMC Breaks Ground on Fab 18 in Southern Taiwan Science park". Press release, Jan 26th. Available at: https://www.tsmc.com/tsmcdotcom/PRListingNewsArchivesAction.do?action=detail&newsid=THGOHITHTH&language=E (accessed 10 July 2020)

TSMC, 2020. "2019 Annual Report" [online]. Annual report, p. 81. Available at: https://www.tsmc.com/download/ir/annualReports/2019/english/ebook/index.html (accessed 10 July 2020)

USTR (United States Trade Representative), 2018. "Findings of the investigation into China's acts, policies, and practices related to technology transfer, intellectual property, and innovation under section 301 of the trade act of 1974". Executive office of the President of the United States.

van Hezewijk, B., 2019. "China Winning the Future of the semiconductor industry?" [online]. SemiWiki.com, Sep 1st. Available at: https://semiwiki.com/general/274582-china-winning-the-future-of-the-semiconductor-industry/ (accessed 10 June 2020)

van Hezewijk, B., 2020. "China's position in the Global Semiconductor Value Chain" [online]. SemiWiki.com, May 25th. Available at: https://semiwiki.com/china/285941-chinas-position-in-the-global-semiconductor-value-chain/ (accessed 10 June 2020)

VerWey, J., 2019. "Chinese Semiconductor Industrial Policy: Past and Present". Journal of International Commerce and Economics, p. 1.

Wade, R., 1990. "Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization". Princeton University Press.

Waldrop, M.M., 2016. "The chips are down for Moore's law". Nature, 530(7589), p. 144-147.

Warren, M.R., 2020. "Bipartisan, Bicameral Bill Will Help Bring Production of Semiconductors, Critical to National Security, Back to U.S.". Press release, Jun 10th. Available at: https://www.warner.senate.gov/public/index.cfm/2020/6/bipartisan-bicameral-bill-will-help-bring-production-of-semiconductors-critical-to-national-security-back-to-u-s (accessed 20 July 2020)

Wei, S., 2020. "SMIC gets $2 billion from China's state-backed funds" [online]. Technode, May 18th. Available at: https://technode.com/2020/05/18/chinas-state-backed-funds-inject-2-billion-into-local-chip-maker-smic/ (accessed July 10 2020)

Wu, D., 2020. "TSMC Scores Subsidies and Picks Site for $12 Billion U.S. Plant" [online]. Bloomberg, Jun 9th. Available (behind paywall) at: https://www.bloomberg.com/news/articles/2020-06-09/tsmc-confident-of-replacing-any-huawei-orders-lost-to-u-s-curbs (accessed 10 Jun 2020)

Yang, Y.K., 2006. "The Taiwanese Notebook Computer Production Network in China: Implication for Upgrading of the Chinese Electronics Industry".

Ye, P., Ernst, T., & Khare, M.V., 2019. "The last silicon transistor: Nanosheet devices could be the final evolutionary step for Moore's Law". IEEE Spectrum, 56(8), p. 30-35.

Yue, C.P. & Lu, T., 2017. "China's Latest Overseas M&A in the Semiconductor Industry". IEEE Solid-State Circuits Magazine, 9(4). p. 8-12.

Zenglein, M.J. & Holzmann, A., 2019. "Evolving Made in China 2025: China's industrial policy in the quest for global tech leadership" [online]. MERICS, Jul 2nd. Available at: https://merics.org/en/report/evolving-made-china-2025 (accessed 10 June 2020).

THE FIFTH WAVE

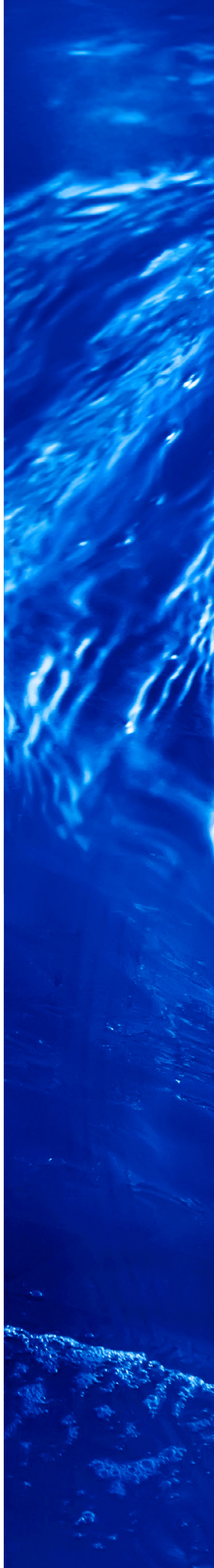
BRIE-ETLA 2019–2023

Over the past decades, information technology has driven societal development and economic growth. The continuing advance of digitalization has enabled individuals and organizations to leverage more and more intelligent tools, leading to disruptive transformations in markets, business models, jobs, and social conventions from time to time. Overall, the field of information and communication technology is one generally characterized by rapid development.

The general focus of the ICT industry and the societal spotlight for the latest expected disruption can quickly move from one technology phenomenon to another. For example, artificial intelligence—while undeniably a hot topic in today's discussion—was hardly heralded as the all-pervasive catalyst for digital transformation three years ago. Similarly, it is likely that during the next three years, new paradigm shifts will occur in the perceived landscape of disruptive technology development.

According to research, the era of machine learning, deep learning and foundation models in the disruption from artificial intelligence is coming to an end. Researchers have not been able to form a consensus on what kinds of transformative developments might be expected to take the spotlight in the post-AI era. Moreover, the pivotal platforms, business models, or intelligent tools essential to those developments have not yet been identified.

The bulk of current societal analysis uses a narrow rear-view perspective by analyzing historic micro- or macroeconomic data. In the case of emerging technologies, however, multi-dimensional interdisciplinary research is required to understand the complex socio-economic mechanisms underlying technology disruptions and how to best navigate businesses and countries in periods of extreme technology-induced turbulence.



ETLA

ISBN 978-951-628-759-4