



Core, intertwined, and ecosystem-specific clusters in platform ecosystems: analyzing similarities in the digital transformation of the automotive, blockchain, financial, insurance and IIoT industry

Tobias Riasanow¹ · Lea Jäntgen¹ · Sebastian Hermes¹ · Markus Böhm¹ · Helmut Krcmar¹

Received: 18 May 2019 / Accepted: 5 February 2020 / Published online: 25 February 2020
© The Author(s) 2020

Abstract

Digital transformation is continuously changing ecosystems, which also forces established companies to re-evaluate their value proposition. However, only transformations of single ecosystems have been studied. Therefore, this work targets to examine the similarities of digital transformation in five platform ecosystems: automotive, blockchain, financial, insurance, and IIoT. For our analysis, we combine the strengths of conceptual modeling using e3 value with a cluster analysis based on text mining to identify similarities in the respective ecosystems. As a result, we identified 15 clusters. Cluster 01 is the core cluster, containing the roles of organizations from all five ecosystems. Cluster 02–05 are intertwined, as they include roles from at least two ecosystems. Clusters 06–15 are ecosystem-specific that only include roles found in one ecosystem. Scholars and practitioners can use these clusters when analyzing or building a new platform ecosystem, or transforming a traditional ecosystem towards a platform ecosystem.

Keywords Platform ecosystem · Ecosystem · Digital transformation · Ecosystem analysis · Similarities

JEL Classification 1.020: Method · 3.060: Business networking · 3.150: Electronic markets · 3.240: Internet of Things · 3.260: Interorganisational Systems

This article is part of the Topical Collection on Electronic markets: evolution and perspectives

Responsible Editor: Hans-Dieter Zimmermann

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12525-020-00407-6>) contains supplementary material, which is available to authorized users.

✉ Tobias Riasanow
tobias.riasanow@tum.de

Lea Jäntgen
lea.jaentgen@tum.de

Sebastian Hermes
sebastian.hermes@tum.de

Markus Böhm
markus.boehm@tum.de

Helmut Krcmar
krcmar@in.tum.de

¹ Technical University of Munich
Boltzmannstrasse 3, Garching 85748, Germany

Motivation

Digital innovations enabled by new technologies fundamentally transform the way organizations interact with each other (Fichman 2014; Yoo et al. 2012). We refer to the organizational transformation to account for these adaptations in value creation through the innovative use of digital technologies as digital transformation (Riasanow et al. 2019; Vial, 2019). Digital transformation also challenges organizations that build only physical products with the need to incorporate digital services as part of their offerings (Vial 2019).

However, most studies regarding digital transformation are primarily concerned with an intra-organizational perspective, such as the transformation of processes, products, and services, organizational structures, or business models (e.g., see Agarwal et al. 2010; Hansen and Sia 2015; Kaltenecker et al. 2015). They should, however, also take an inter-organizational perspective into account (Jacobides et al. 2018; Puschmann 2017), particularly since digital transformation may substantially influence

inter-organizational partnerships in ecosystems when value is co-created among multiple stakeholders (Sarker et al. 2012).

Increasingly a structured analysis of ecosystems has gained attraction among scholars in information systems, management, and organization science (Adner 2006; Autio and Thomas 2014; Jacobides et al. 2018; Tiwana 2015). To visualize and analyze ecosystems, cluster analysis (Basole 2009; Basole et al. 2018), ecosystem-as-a-structure (Adner 2017), or conceptual modeling have been used (Riasanow et al. 2017). However, the analyzed ecosystems only concern single industries (see Adner 2017). Therefore, we still lack a structured analysis of the similarities of digital transformation in multiple ecosystems (Jacobides et al. 2018; Vial 2019). Hence, when analyzing, building, or transforming ecosystems, scholars and decision-makers often disregard their inherent interconnectivity.

Therefore, this study analyzes five platform ecosystems, which we use to answer the following research question: *What are the similarities in platform ecosystems in a digital transformation?* For our analysis, we suggest a new method to identify and analyze similarities among several platform ecosystems by combining the strengths of conceptual modeling and cluster analysis.

The paper is organized as follows. First, we provide a brief survey of related work on platform ecosystems and existing methods to model and analyze ecosystems. Second, we describe our method for identifying similarities in a two-phased process. First, we draw on Crunchbase data to drive roles for organizations in one ecosystem and use the conceptual modeling technique e3 value for visualization (phase 1). Second, we perform a cluster analysis based on the Crunchbase data of all organizations using text mining to identify similarities among the five ecosystems (phase 2). Based on that analysis, we identified 15 clusters. Cluster 01 “Cloud and On-Premise Providers, Cyber Security Providers” is the core cluster, as it contains roles of organizations of all five ecosystems, with some of them (cloud application, cloud platform, and cloud infrastructure provider) found in all ecosystems. Notably, based on the similarity of the organization descriptions, roles like disruptive hardware, IIoT solution, or digital identity providers can be found in the core cluster, too. Cluster 02–05 “Digital Financial Services”, “OEMs and IIoT Solutions”, “Data Prediction and Monitoring”, and “Brokers and Agents” are intertwined clusters, as they include roles from at least two ecosystems. Cluster 06–15 are ecosystem-specific, which, therefore, only contain roles found in one of the five ecosystems. We conclude with a discussion of our results, the limitations of this work, and suggest avenues for future research. Scholars can use these core roles when analyzing platform ecosystems, and practitioners can use them when building or transforming a traditional ecosystem towards a platform ecosystem.

Related work

Digital platforms that have the capacity to combine and deploy innovative technologies create new business models that fundamentally transform the way business is done (Hein et al. 2019; Lucas and Goh 2009; Tiwana 2015). We refer to the organizational transformation to account for these adaptations in value creation through the innovative use of digital technologies as digital transformation (Riasanow et al. 2019; Vial 2019).

Yet, studies about digital transformation initiatives are primarily concerned with an intra-organizational perspective on transformations, such as processes, products, and services, organizational structures, or the business model (see, e.g., Kaltenecker et al. 2015; Karimi and Walter 2015). However, digital transformation initiatives substantially influence inter-organizational partnerships, particularly in platform ecosystems, where value is co-created among multiple stakeholders (Ceccagnoli et al. 2014; Sarker et al. 2012).

Platform ecosystems

Three terminologies for ecosystems are most commonly used in IS, management, and organization science research, which also divide the field into three broad streams, as found by Jacobides et al. (2018): “business ecosystems”, “innovation ecosystems” and “platform ecosystems”. The three streams differ in their focus of the research but share the common understanding of ecosystems as a group of interdependent, but loosely coupled firms. In a hierarchical sense, the term “business ecosystem” can be seen as the root, with “innovation ecosystems” and “platform ecosystems” derived thereafter.

Moore (1993) defined business ecosystems as companies with “co-evolv[ing] capabilities around a new innovation” in a cooperative and at the same time competitive way. This broader definition has since been widely maintained, with Teece (2007) defining business ecosystems as “the community of organizations, institutions, and individuals that impact the enterprise and the enterprise’s customers and suppliers”. There is still uncertainty as to where exactly the boundary has to be drawn that separates the entities within from that outside of a specific ecosystem (Weber and Hine 2015). However, this definition also entails that a business ecosystem is constituted relative to a particular firm – with different firms that are not operating in the identical ecosystem even when they are offering similar services or products (Jacobides et al. 2018).

In some of the definitions of business ecosystems, the term “platform” is already mentioned, as in the conceptualization of Autio and Thomas (2014). This already indicates how closely the idea of a platform is related to ecosystems. Ecosystems are the more generic concept, of which platform ecosystems are one typical instantiation: Many ecosystems, such as the Apple

iOS ecosystem, have at their core a platform that structures and orchestrates the complementors and partners (Altman and Tuschman 2017; Dattée et al. 2018). The term “platform ecosystem” is likely to be most conventional in IS, shaped, for example, by the work of Tiwana on ecosystems around software platforms (Tiwana 2014; Tiwana et al. 2010). The designation “platform” originates from the product development or engineering disciplines (Kirshnan and Ulrich 2001; Simpson et al. 2001), and has since enjoyed similar popularity as the term “ecosystem”, with further adoption in fields such as (industrial) economics (Evans 2003; Rochet and Tirole 2003).

In management literature, it has also gained significant momentum in the meantime (Boudreau and Lakhani 2009; Cusumano and Gawer 2002; Hagi 2014), especially in researching the mechanisms of two- or multiple-sided markets. In the review by Thomas, Autio, and Gann (2014), the literature on platforms in management was consolidated, with “platform ecosystems” as one of the major streams.

Approaches to ecosystems analysis

Platform ecosystems with actors that have unique, supermodular, or non-generic complementarities require the creation of a specific structure of relationships to create value (Jacobides et al. 2018). Different methods exist to first visualize and second analyze ecosystems, such as conceptual, heuristic, mathematical or ontological methods, as well as cluster analysis (Arreola Gonzalez et al. 2019; Basole et al. 2018).

We focus our attention on conceptual modeling using the e3 value methodology (Gordijn and Akkermans 2003), and cluster analysis based on text mining. This design choice is supported by two considerations. First, the topic of value ecosystem modeling and analysis is simply too huge to be acceptably covered in a single survey paper if prior work is to be recognized in any serious fashion. Second, we focus on e3 value for conceptual modeling based on its suitability to comprehensively visualize large ecosystems (e.g., Böhm et al. 2010), and cluster analysis using text mining as it allows to objectively analyze large ecosystems based on organizational descriptions (see Basole et al. 2018).

Conceptual modeling

First, ecosystem-as-a-structure is the easiest conceptual method to visualize and analyze ecosystems. Ecosystem-as-a-structure models every participant relevant to an organization of an ecosystem (Adner 2017). Since its level of abstraction is typically on an individual organization level, this is the easiest way to visualize and analyze the ecosystem around a single organization or digital platform. This approach is also used to identify the value creation of organizations in ecosystems (Urmetzer et al. 2018). However, if an ecosystem is large,

ecosystem-as-a-structure models become very large and would require more abstraction.

The abstraction of similar organizations to roles can be modeled with e3 value, another conceptual modeling technique (Gordijn and Akkermans 2003; Gordijn et al. 2006). Therefore, the e3 value methodology is more suitable to model large ecosystems consisting of a multitude of different actors. For example, the generic cloud computing network consists of infrastructure providers, platform providers, application providers, a market platform, aggregator, integrator, consultant, and consumers (Böhm et al. 2010). For the aggregation to generic roles, Böhm et al. (2010) used a qualitative content coding technique (Mayring 2010). There, the role application provider contains services such as Dropbox, Microsoft Office 365, or Salesforce CRM, which are all typical instantiations of this role. This reduction of complexity helps to communicate the ecosystem and changes occurring in it more easily. Henceforth, the level of abstraction varies from an organizational level in ecosystem-as-a-structure to a role-based level, which in this method can be the aggregation of different organizations in an ecosystem. However, conceptual modeling is time-consuming and follows a manual process.

Extending the approach of Böhm et al. (2010) to use the e3 value method to visualize the ecosystem, Riasanow et al. (2017) validated their generic automotive ecosystem with expert interviews in the respective ecosystem. This step is conducted to verify the robustness of the developed ecosystem, as the experts discuss and verify the identified roles and their relationships in the ecosystems. Riasanow et al. (2017) identified that the traditional automotive ecosystem is blurring due to the impact of cloud-based roles, which allow the development of digital platforms or further value-adding services, which are partially extending or substituting the value creation of OEMs.

Cluster analysis

The quantitative analysis of ecosystems relies on text mining to cluster organizations regarding their similarity (Basole 2009; Basole and Karla 2011). Thereby, computation power is used to detect patterns and clusters in a much faster way using graph theory algorithms (Basole and Karla 2011). Since organizational data is mostly available in an unstructured form, text mining techniques are leveraged together with graph theory algorithms to detect ecosystems and structural characteristics among this unstructured data (Basole et al. 2018). This method can be used to cluster organizations based on their similarity in description or value creation (Basole et al. 2018).

One of the central advantages besides relying on computation power for the analysis is the objectivity of this methodology. Whereas conceptual modeling relies on the subjective coding, the clusters are mathematically derived based on the

similarity of organizational descriptions, for example. In contrast, this method is limited to the accuracy of the organizational descriptions.

Due to this clustering, Basole et al. (2018) identified a tremendous growth of Fintechs in the last decade, however, the rate of new emerging startups is getting slower, partially due to a maturing of the industry. Second, they identified an increasingly global footprint of the Fintech organizations in their analyzed ecosystem. Third, a core set of Fintech ecosystem players could be identified, and a number of peripheral actors (Basole et al. 2018). This may be due to the financial power of large incumbents such as American Express, JP Morgan, or Bank of America, which are acquirers of Fintech companies that occupy differing structural positions in the ecosystem (Basole et al. 2018). Ultimately, the clustering identified six core market segments (Basole et al. 2018).

Research approach

Our suggested approach aims at combining conceptual modeling with cluster analysis in order to identify similarities among multiple ecosystems. In the first phase, we follow the approach of Böhm et al. (2010) and use e3 value to model multiple ecosystems, and Riasanow et al. (2017) to use Crunchbase to derive organizational data to code generic roles and value streams between them, which we require to model an ecosystem. Furthermore, expert interviews are conducted to validate the ecosystems (Myers and Newman 2007). In the second phase, we adapted the method of Basole et al. (2018) and use the Crunchbase descriptions of all organizations from five ecosystems to perform a cluster analysis to detect similarities among them. Figure 1 provides an overview of the steps of our research approach.

We do not choose a purely quantitative approach, because as we compared our results of the manual modeling of every ecosystem with the results of using the below proposed quantitative approach to assign organizations to roles and they strongly differed. The comparison shows that the quantitative approach is on average only able to predict 53% of the organizations correctly, see Appendix 7. Hence, we choose the proposed combined approach, to profit from the accuracy of manual coding and the objectivity of an unbiased similarity calculation using text mining. In other words, the manually derived roles add additional information to the calculated clusters, which eases interpretation and increases explanatory power.

For our cluster analysis to detect similarities, we combine the following ecosystems: automotive, blockchain, financial and insurance. Further, we will use data from the Industrial Internet of Things (IIoT) organizations to model a new ecosystem, which we use to introduce our approach for the conceptual modeling using e3 value in the subsequent chapter.

We chose the respective five ecosystems based on a theoretical sampling strategy (Eisenhardt 1989; Eisenhardt and Graebner 2007; Yin 2014). We decided for the automotive, financial, and insurance ecosystem to analyze platform ecosystems that are currently substantially transforming due to innovative digital technologies (Basole et al. 2018; Piccinini et al. 2015; Puschmann 2017). We chose the three ecosystems due to their varying maturity of the digital transformation, with the financial ecosystem being most mature due the plethora of emerging organizations, which are breaking up and recombining the value creation of incumbent organizations (Basole et al. 2018; Westerman et al. 2011), and the automotive ecosystem being less mature, mainly due to its asset-heavy products (Piccinini et al. 2015; Riasanow et al. 2017). In the next step, we focused on ecosystems based on innovative technologies with the capabilities to transform or substitute products and services of incumbents in other ecosystems and chose the blockchain and IIoT ecosystem. For example, cryptocurrencies allow payments without trusted intermediaries such as banks (Tapscott and Tapscott 2016), and IIoT is an enabler for location-based or pay-per-use insurances in the car (Desyllas and Sako 2013; Greineder et al. 2019). In other words, both technologies serve as baseline for the digital transformation of the automotive, financial, and insurance ecosystem.

As next, we will briefly explain all of the steps of phase 1, the conceptual modeling, and provide the input, procedure, and results of each step for the IIoT ecosystem. The data of the other ecosystems are based on prior work (Greineder et al. 2019; Riasanow et al. 2018a; Riasanow et al. 2018b; Riasanow et al. 2017), which can be found in Appendix 1–5. Afterward, we present the steps of phase 2.

Inductive coding and conceptual modeling

We use Crunchbase, a socially curated directory of organizations, people, and investors, in order to extract organizational data to model our ecosystems. Following Basole et al. (2018), due to a large number of entries, Crunchbase data is suitable to model ecosystems. We use the provided information about these organizations to derive roles and value streams. Crunchbase possesses a comprehensive database of traditional companies, or incumbents, and startups (Marra et al. 2015), including a description of organizations' value propositions. Additionally, start-ups at all funding stages are included in the database, which enables researchers to capture new business model innovations in emerging markets (Marra et al. 2015; Yu and Perotti 2015). The information reported in the database consists of the company size class, company descriptions, its location, its primary role (firms, group, or investor), its status (operating, acquired, IPO, or closed), its founding date, and the dates on which the record was created and updated (Basole et al. 2018). All additions and edits in the Crunchbase data

undergo a verification process before they are released online. Crunchbase, therefore, allows the capture of established and emerging organizations related to an ecosystem and related technologies.¹

Additionally, other databases for specific ecosystems can be used as well. As an example, we used the IoT One database,² which provides a comprehensive overview of existing IIoT platforms and solution providers.

Data extraction

As input for the extraction of the company data for our IIoT ecosystem, we used the following search terms: “IIoT”, “Industrial IoT”, “Industrial Internet of Things”, and “Industry 4.0” for the IIoT ecosystem. On May 29th, 2018 we extracted the data of 308 organizations from Crunchbase and an additional 200 organizations from the IoT One database “Top 200 IIoT companies”. Since top companies in the field of IoT were still missing according to CB Insights, we added 115 organizations from their “Top 125 IoT Startup” list. As the output of this step, we extracted data from 633 organizations. As an example of the extracted data, Appendix 10 shows the data of Alleantia,³ one of the extracted organizations.

Data screening

Using the extracted data from the 633 organizations, we excluded all organizations that were marked as “closed” or “acquired” in the operating status since our goal is to evaluate the current situation of an ecosystem.

For example, Lumenetec, an award-winning software, sensor, and analytics innovator providing performance optimization, big data analytics, and risk management, was closed in 2015. Linear AMS, a provider of tooling for injection molds, compression molds, and tooling required for processing of thermoplastic resins, was acquired by Moog in 2015. Additionally, we excluded all organizations that did not include a website or description provided by Crunchbase, which amounted to 11 organizations. Screening the data, we also found companies that had no relationship to the IIoT industry. Hence, we shortened the data set by a further 21 companies. For the IIoT ecosystem, 601 organizations remained in total for further analysis.

¹ For data gathering we used a Crunchbase Premium account, since the free account limits the use (and amount of) available company data

² <https://www.iotone.com>

³ The data can be also found online: <https://www.crunchbase.com/organization/alleantia>

Coding of roles and value streams

To identify roles and value streams in an ecosystem, we use structured content analysis, including an inductive category development based on Mayring (2010) and Miles and Huberman (1994). First, one of the coders used the organizational descriptions derived from Crunchbase to develop codes for the generic roles. For example, the generic role IIoT Solution Provider is connected to the terms: solution, scalable, data collection, from wire to cloud, software, interoperable, connect devices (e.g., sensors), or digital twin, see Table 1.

Next, inspired by the codes and the organizational descriptions from Crunchbase, descriptions for the generic roles were developed, for example, see Table 2.

Afterward, the organizational descriptions and the descriptions of the generic roles were given to a second rater, who coded the organizations to the generic roles independently. Both raters compared and discussed their coding for calibration purposes. To measure the intercoder reliability, we calculated Krippendorff’s (2004) Alpha. The results indicated an Alpha of 0.83, reflecting acceptable intercoder reliability (Krippendorff 2004).

All authors confirmed the final coding of each organization and discussed discrepancies, which helped to eliminate individual disparities (Bullock and Tubbs 1990).

The same approach is used for the identification of the value streams, but we combined the Crunchbase information with secondary publicly available information from company websites, reports, press articles, or annual reports. Coding the remaining 601 organizations of the IIoT ecosystem, we identified a set of 14 generic roles. Table 2 shows three of the 14 derived roles with example organizations, the other can be found in Appendix 6.

Visualization and validation

In the next step, we use the e^3 -value method to visualize the IIoT ecosystem based on the identified roles and the value streams between these (Böhm et al. 2010; Gordijn and Akkermans 2003). The e^3 -value method is a business modeling methodology to elicit, analyze, and evaluate business ideas from an ecosystem perspective. It is used to assess the economic sustainability of an ecosystem by modeling the exchange of economic value between actors (Gordijn and Akkermans 2003). The visualized IIoT ecosystem can be seen in Fig. 2.

To validate the ecosystem, interviews with industry experts were conducted. We used a semi-structured technique (Myers and Newman 2007). Each of the chosen experts demonstrates a wealth of experience in the respective ecosystem, as well as with digital technologies. We use experts in a leading strategic position or information technology-related function (Goldberg et al. 2016) who have privileged access to

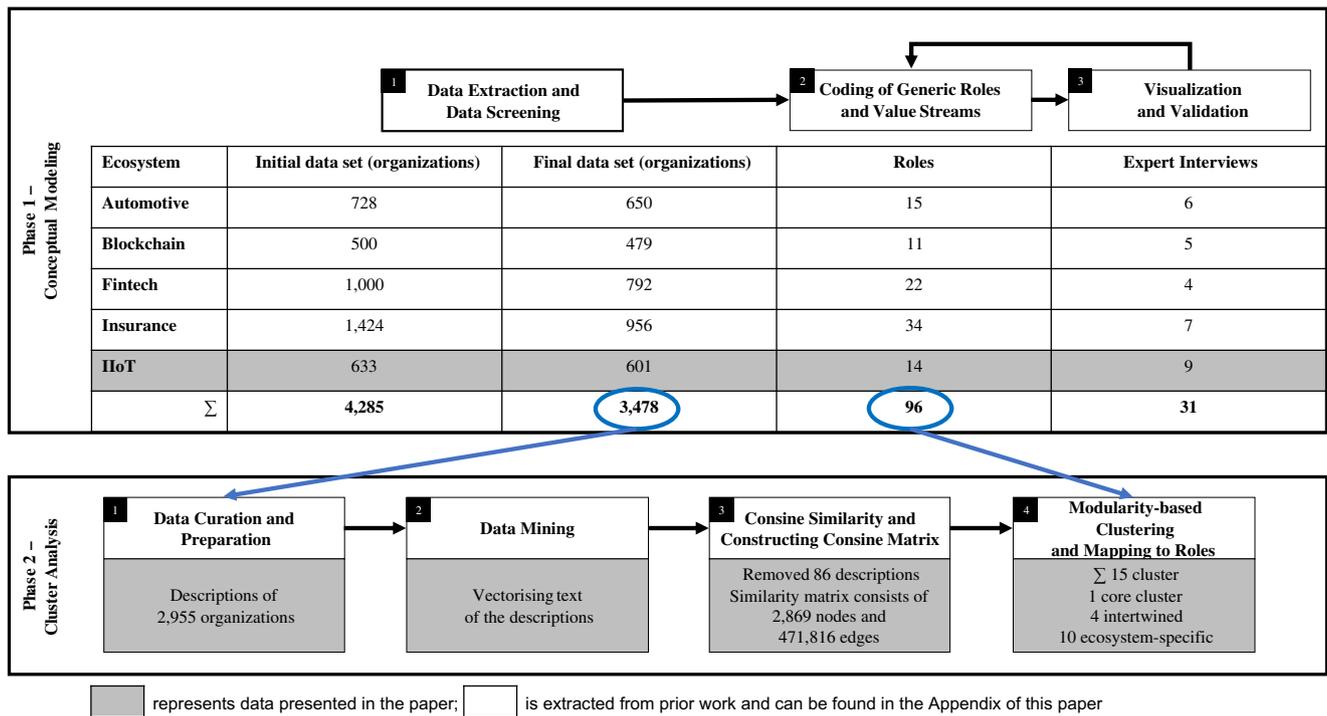


Fig. 1 Overview of our Research Approach

information and knowledge on the subject (Bogner et al. 2009). This allows us to draw on different practice-based insights from various companies and long-time market experience. We recorded all the interviews and transcribed them. During the interviews, we discuss the roles and value streams of the proposed modeled ecosystem with the experts to validate the modeling.

For the IIoT ecosystem, we conducted nine interviews with experts from the manufacturing industry to validate the generic ecosystem. Interviewees were the head of department of “Industry 4.0”, a project leader for Industry 4.0, a partner for production, a partner for Industry 4.0, a Head of Industrial Research and Innovation, a Process Manager, a consultant for Industry 4.0 in manufacturing companies, and two partners of leading consulting companies with a long history in the manufacturing business. Each of the experts has substantial experience in the manufacturing industry and IIoT technologies. We conducted the interviews between August and September 2018. The interviews took 38 minutes on average. Our

interviewees confirmed the identified roles and did not suggest new roles for the IIoT ecosystem.

The visualization and information regarding the conducted expert interviews for the remaining four ecosystems can be found in Appendix 1–5.

Clustering based on similarity measures using text mining

After coding every ecosystem iteratively, we followed Basole et al. (2018) to detect similarities in the five ecosystems—automotive, blockchain, financial, insurance, and IIoT.

Data curation and preparation

We use the previously extracted Crunchbase data of all five ecosystems to perform the text-mining based cluster analysis. The Crunchbase description field contains textual and unstructured data on organizations. We started with the 3478 previously coded organizations. We had to drop

Table 1 Roles and descriptions of the actors in the IIoT ecosystem

Organization	Crunchbase description (extract)	Coded role
Alleantia	(...) Alleantia provides the most <u>scalable</u> and cost-effective <u>Industrial IoT solution</u> in the market for <u>data collection</u> , <u>information standardization</u> and <u>distribution from wire to cloud</u> , to implement pervasive machine data access and comprehensive information sharing within the enterprise and across its extended supply chain. Alleantia software connects in a few seconds <u>any device</u> — from complex machinery to simpler sensors — to <u>create a complete and interoperable digital twin</u> (...).	IIoT Solution Provider

Table 2 Roles and descriptions of the actors in the IIoT ecosystem (excerpt)

Role and Description	Example(s)
<p>Sensor and Connectivity Providers offer connectivity for entities used by the manufacturer/OEM. This includes sensors to connect machines already in use, or sensors for new machines, or robots that allow connectivity to the manufacturing process. The sensors can be used to collect data, which is an essential prerequisite for IIoT applications, or offer wireless network connectivity for IoT devices, and thus, a connection to the cloud.</p>	Alien Technology, Libelium, Verizon, Acent Systems
<p>The IIoT Hardware Provider offers connected hardware, technology, or machines used for the digital factory. Members of this role typically possess proven expertise in the engineering required for the hardware (and often for the complementary software).</p>	Kuka, EOS, Rethink Robotics
<p>The IIoT Solution Provider offers complete solutions of hardware and software to manufacturers. Therefore, he/she uses the offerings of other providers like the Industrial IoT Hardware Provider, Sensor & Connectivity Provider, Fog & Edge Intelligence Provider, and the Added Value Service Provider or partly fulfills those roles himself. Compared to the Hardware or Sensor provider, he/she is not only offering the product but rather a complete solution to perform a service.</p>	Alleantia, Konux, Xometry, KAESER

522 of the organizations since they were no longer available or were not imported via Crunchbase (e.g. OneIoT Database). After data curation, 2955 data sets were used for further steps. Data sets for each organization include the organization name, the assigned role, the ecosystem and the organization description extracted from Crunchbase.

Data mining

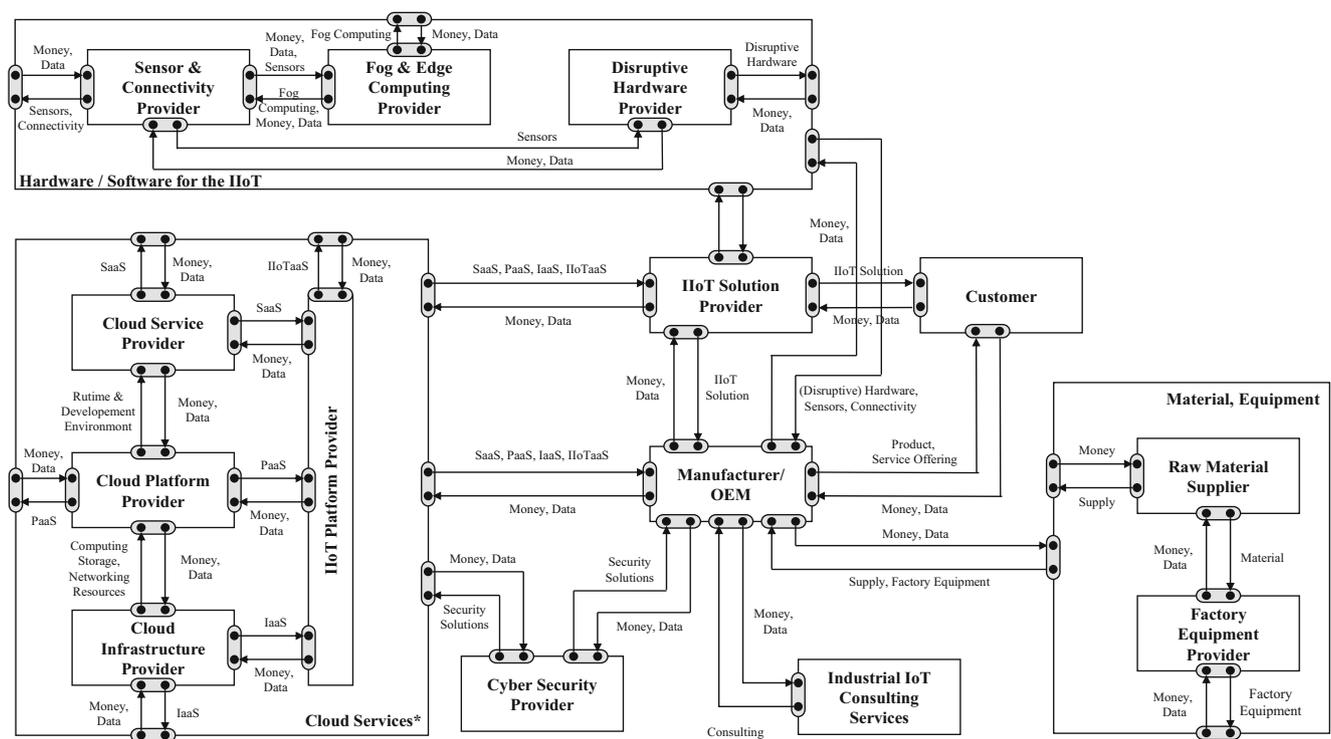
In this step, we clustered the remaining 2955 organizational data sets.

In these data sets, the description of each organization is provided as unstructured text and has to be converted to a

computer-readable form. We used text analytics to convert the text into vectors of words. First, we removed all stop words. This condenses the text by only keeping words with a real meaning (Basole et al. 2018).

Next, we used Porter’s suffix-stripping algorithm for stemming the words (Porter 1980). Stemming is a normalizing procedure since several variations of words carry the same meaning (Metzler et al. 2007). We follow Basole et al. (2018) and use term frequency-inverse document frequency (TF-IDF) as a well-established weighting method for vectorizing text.

Since not all words in a text are equally important, more frequently appearing words can be seen as carrying more



*IIoT Platform can also be installed on premise

Fig. 2 Generic IIoT ecosystem

information about the text. However, if the term appears frequently across all texts, it loses its distinguishing power. Therefore, TF-IDF normalizes the frequency of words in a text with the rarity it appears. Described by Ramos (2003), we use TF-IDF, which formally assigns weights to words as a combination of a local measure on a description basis and a global measure on all descriptions combined. This ensures that ecosystem-specific words, such as “automotive”, “finance” or “insurance” are assigned lower weights since these words do not carry information about value propositions of the organization. For the implementation, we use the sklearn package of Python.⁴

Computation of cosine similarity and constructing the similarity matrix

After data curation, we identify similarities between the organizational descriptions. To do so, we use similarity measures, which are tools for calculating the degree of similarity between two objects, in our case vectors built using TF-IDF.

Following Basole et al. (2018), we use the cosine similarity for calculating the similarity between vectors of organization descriptions. Cosine similarity quantifies similarity by the cosine of the angle between two vectors. The cosine similarity is calculated pairwise between the organizations and represented for all organizations as Matrix A of shape $N \times N$, where N is the number of organizations in the analysis.

The results range from 0 to 1, where 1 represents the equality of the texts. We assume that organizations coded in the same role should be treated as identical since we already know the connection between them. We, therefore, change the similarity measure of these organizations to 1. By doing this, we enhance the data with the manual work of the previous step.

Matrix A can be seen as an adjacency matrix for a graph representing organizations as nodes and similarities as edge weights between them. For constructing the graph, we take the lower triangular suggest of A and exclude the diagonal as well. This way, edges between organizations are only considered once and similarities between the same organizations are excluded. In addition, following Basole et al. (2018), we removed 86 organizations that are not similar to any other organization and therefore represent nodes in the graph that are not connected.

We use the NetworkX package in python to construct the graph.⁵ The resulting graph consists of 2869 nodes and 472,816 edges.

Modularity-based clustering

After constructing the graph based on similarity measures, we target to identify clusters.

We identify clusters in a graph or network based on the computation of modularity (Blondel et al. 2008). Following Fortunato and Barthelemy (2007), modularity measures the strength of the division of a graph into subgroups. Modularity is maximized if densely connected nodes are clustered together and not as frequently linked nodes are split into different clusters (Newman 2006).

Following Basole et al. (2018), we choose Louvain’s modularity-based clustering algorithm due to its good performance in large graphs with a good quality of clusters. As a result, we identify 15 clusters containing organizations from different roles from the five input ecosystems (automotive, blockchain, financial, insurance, IIoT). The clusters can be found in Appendix 8.

In order to analyze the clusters, we aggregate all organizations included in the clusters to the generic roles as shown in the previous qualitative steps. E.g., BMW and Daimler are aggregated to the generic role OEM.

Cluster visualization and latent semantic analysis for topic detection

We use Gephi⁶ to visualize the created graph of organizations and clusters. Visualizations are important for human understanding and enable us to better understand the clusters (Basole et al. 2016).

Since there is no best solution for representing data, the choice should be guided by the nature of the data and the question that needs to be answered (Basole et al. 2018). In our case, a visual representation of the graph showing interconnectivity between organizations and clusters has to be taken into account. We follow Basole et al. (2018) and apply the noverlap algorithm in order to prevent nodes from overlapping. We color the individual nodes according to their ecosystem and mark the clusters with a squared frame in order to improve readability. We also use curved edges in order to make the graph more appealing. We color edges based on the ecosystem of the source node, see Fig. 3.

In order to gain more insight into the clusters, we perform latent semantic analysis (LSA) (Deerwester et al. 1990) on individual clusters in order to extract keywords and topics from the descriptions or organizations in clusters. LSA is an unsupervised text analytics algorithm using statistical measures in order to find a hidden meaning of word usage. Also known as Latent Semantic Index (LSI), LSA is used in natural language processing for feature extraction and information retrieval.

In the first step, we remove all ecosystem-specific words and preprocess the data by removing stop words and stemming the words (see above). We only remove the ecosystem-specific words here for the topic detection (we did not remove

⁴ <https://scikit-learn.org>

⁵ <https://networkx.github.io/>

⁶ <https://gephi.org/>

them in the clustering before), as we seek to identify similarities in the ecosystems. Hence, removing ecosystem-specific terms ensures that the cluster descriptions only include words connected to the value proposition of the organizations.

We then represent the organization descriptions of each cluster as a term-document matrix M . Columns in the term-document matrix represent organizations and rows of all words occurring in these descriptions. The values in the term-document matrix represent the importance of the word; we use the above-mentioned TF-IDF method for quantifying the importance of a word in each description since a weighted matrix has proven to resolve in better results (Dumais 1991). Next singular value decomposition (SVD) is used on the term-document matrix in order to break down the matrix into topics. Through this process, latent meaning, noise reduction, high-order co-occurrence, and sparsity reduction can be achieved (Turney and Pantel 2010). We then use the cosine similarity measure in the reduced term-document matrix vector space for measuring the closeness between the words.

For this method, we need to make a choice about how many dimensions should be removed, or in other words how many topics do we want for any given cluster. We use a coherence score in order to identify the optimal number of topics we want to retrieve but set the maximum to 10 since we want an interpretable output. For the calculation, we use the Gensim coherence model implemented in Python based on Röder, Roth, and Hinneburg (2015). The coherence score is commonly used for evaluating topic models in which a higher coherence score indicates a better model. For each cluster and each number of topics between 1 and 10, we build an LSA model using the Gensim package and evaluate the coherence score. We then choose the number of topics for each cluster that maximizes the coherence score. Clusters can, therefore, have a different number of topics assigned to them.

Cluster analysis

Following the described steps, we identified 15 clusters for the 2955 organizations originating from the five ecosystems (blockchain, financial industry, automotive, insurance industry, IIoT). As a result, we mapped the organizations in the 15 clusters to their generic roles assigned in phase 1. Hence, we can use the generic roles to discuss the various value propositions found in the clusters. The LSA topics are guidelines for naming the clusters and give reference for the interpretation. Figure 3 shows the visualization of the clusters.

Towards our search for similarities in the digital transformation of platform ecosystems, we first focus on 5 of the 15 clusters, as they contain organizations from more than one ecosystem. Cluster 01 is the “core” cluster, as it contains roles from all five ecosystems, with some of them found in all ecosystems. Cluster 02–05 are “intertwined”, as they include

roles from at least two ecosystems, marked black in Figure 03. Cluster 06–15 are “ecosystem-specific” since they contain organizations from only one ecosystem, marked gray in Fig. 3.

For each of the clusters, we briefly provide the name, a short description, the extracted topics, the generic roles these organizations were assigned to in phase 1, and their affiliated ecosystems. Finally, we discuss the identified connection between the identified clusters. The assigned roles for each cluster, as well as the calculation results for the coherence value of the keyword extraction, for each of the clusters, can be found in Appendix 8. Appendix 9 shows the result of the sensitivity analysis.

Core cluster

The first cluster, **Cloud and On Premise Infrastructure Provider, Cyber Security Provider**, contains organizations, and therefore, generic roles of all five ecosystems. Using the LSA Algorithm, we identified two topics from which we extract the first three keywords: 1) cloud, application, data, and 2) enterprise, web computing. These topics suggest that containing organizations have an overall cloud and web-computing topic.

Mapping the organizations of this cluster to their generic roles (in phase 1), this cluster contains organizations from three roles that we found in all ecosystems: cloud infrastructure, platform, or application providers, such as so-called “hyperscalers” Microsoft Azure, AWS, Google Cloud, or the Alibaba Cloud Platform. Based on the scalability of cloud infrastructure, such as AWS’ Elastic Compute Cloud, or the Microsoft Azure Cloud Platform, or cloud applications, such as Dropbox, or Facebook, the organizations connected to these roles can be seen as the core of all digital platform ecosystems. Precisely because these generic roles build industry-independent services, they can be used in so many ecosystems. As many of these services, such as offered by AWS, are interpreted as commodities, these organizations follow a cost leadership strategy. In contrast, other innovative companies have developed new services based on these commodities, which now extend or substitute the value creation of incumbents (Vial 2019; Westerman et al. 2014). This is particularly visible in our financial ecosystem, where a large number of start-ups offer new payment or financing options such as crowd funding or crowd lending (Basole et al. 2018; Puschmann 2017). Also, pay-per-use insurances in the insurance ecosystem or mobility data providers in the automotive ecosystem are largely building on cloud infrastructure, platforms, or services (Greineder et al. 2019).

Additionally, this cluster contains organizations in generic roles connected to data protection and security, such as the Cyber Security Provider, Digital Identity Provider, Fraud Prevention Provider, and Regulatory Authorities. Security-related organizations and regulatory authorities

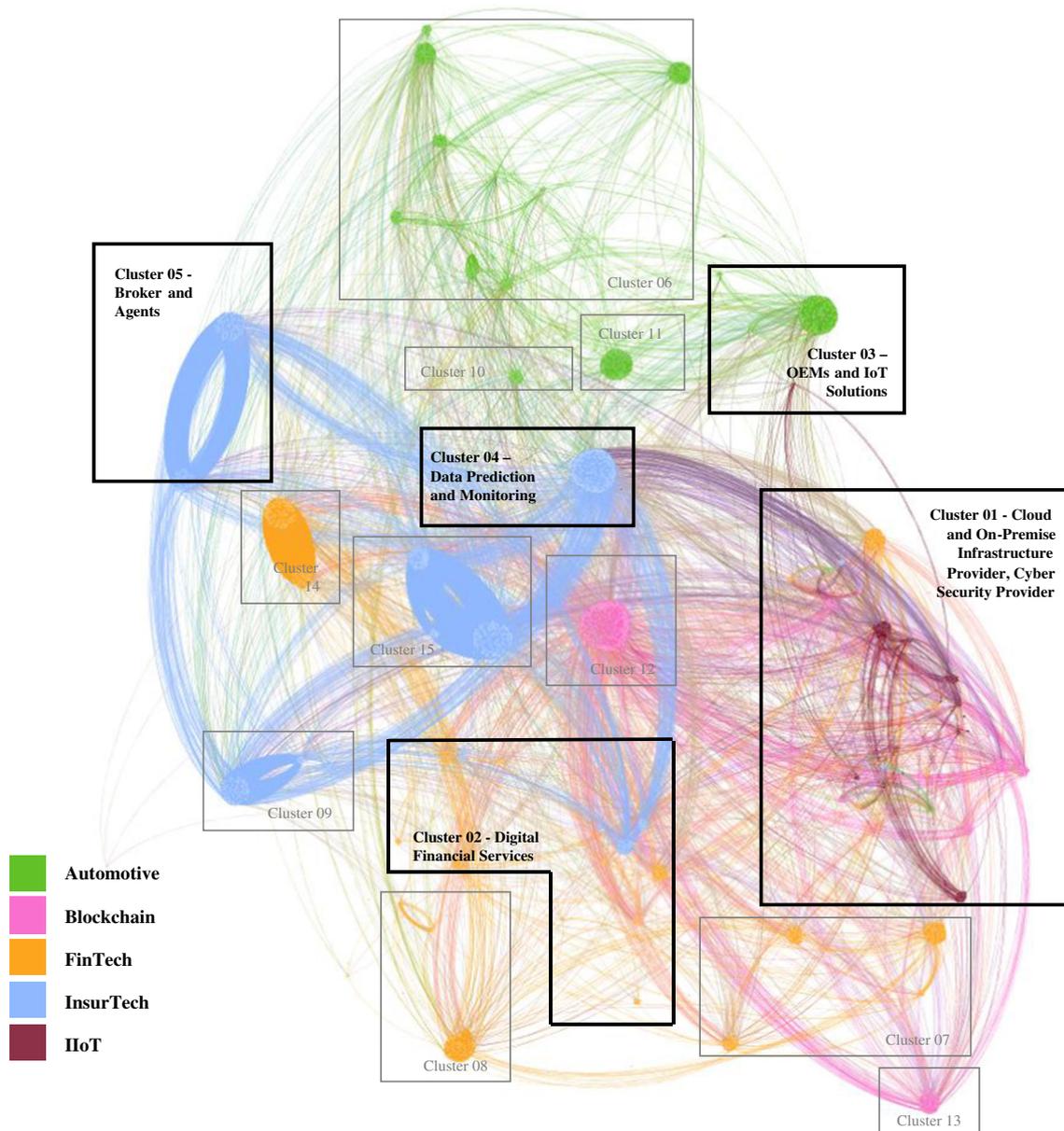


Fig. 3 Visualization of clusters

are a core component of the analyzed platform ecosystems. Notably, digital identity providers are similar to cloud services, as many cloud services require identification prior to their usage. Additionally, digital platforms are particularly confronted with regulatory issues such as the expensive insurance license that InsurTech Element had to acquire to provide property, accident and liability insurances as a service.

However, based on the similarity of their organizational descriptions, we also find disruptive technology roles, such as Mobileye, an organization that builds a camera-based system that serves as a basis for autonomous driving, in this cluster. Also smaller roles, such as Mining Pool, SaaS, and the Blockchain Community can be found here.

Figure 3 shows that this cluster is very connected, particularly with Clusters 04, 12, and 13. Clusters 12 and 13 represent the application and platform providers of the blockchain industry, which typically have a high interaction with other cloud-based services. We find this convincing, particularly since mobile, cloud-based services contributed greatly to digital transformation in various industries, e.g., cloud-based Fintechs like PayPal or TransferWise in the financial ecosystem.

Intertwined clusters

The second cluster, **Digital Financial Services**, contains organizations from two ecosystems. Using LSA to check for

topics, we found one topic from which we extract the first three keywords: 1) money, currency, digital. As these keywords suggest, the cluster only contains organizations from the financial and insurance industry.

Overall, this cluster consists of generic roles such as Robo-Advice / Portfolio Management, Multi-banking Aggregators, blockchain-based Smart Contracts, Cryptocurrency Exchanges, Instant or P2P Insurances, Risk Management, Saving Accounts and others. Hence, the roles assigned to this cluster reveals the underlying interconnectivity between the financial and insurance ecosystem also found by Basole et al. (2018). Many of these services are building on cloud services found in the first cluster. This may be due to the financial ecosystem being one of the first industries that started a digital transformation after the broad availability of cloud services (Puschmann 2017).

Due to the integration in digital platforms like Uber, Netflix, Airbnb, digital financial services, most prominently PayPal, have significantly contributed to the success of these platforms. In the case of Uber, customers, for example, do not have to use cash or credit cards to pay for their ride. Instead, the payment does not require any further interaction as Uber charges customers automatically, which is more convenient for the customer than prior payment methods.

Figure 3 also shows a sub-cluster in Cluster 02 is very connected to Cluster 04 and 05. This sub-cluster contains organizational roles, such as Instant Insurance, Smart Contracts Blockchain, and Cryptocurrency Wallet. Examples for this cluster include the cryptocurrency wallet MyEther-Wallet, or Tröv as Instant Insurance. However, organizations that conduct analytics in the financial and insurance ecosystem are also included in this cluster. Other than this connection, Cluster 02 is not very connected to the rest of the graph.

The third cluster, **OEMs and IoT Solutions**, also contains organizations from two ecosystems: the automotive and IIoT ecosystems. Using the LSA Algorithm, we identify one topic with the following keywords: 1) system, manufacture, develop.

Most of the assigned generic roles in this cluster mostly refer to manufacturing organizations. One example is PINpoint, which offers solutions to manufacturers to cope with Industry 4.0, or Mobileye, which develops camera-based services that serve as the basis of autonomous driving. Hence, these roles are considered the innovative ones in the respective industries. Therefore, these roles are the attackers of the value creation of the traditional roles in the respective ecosystems. Notwithstanding, these organizations cut some market shares of the value creation of traditional manufacturers, on the other side they force them to stay innovative as well.

Cluster 03 also contains some classical roles, such as car manufacturers like BMW and Daimler. As mentioned, these are the organizations that seem to be threatened by digital

transformation in the respective ecosystems (Ceccagnoli et al. 2014; Remané et al. 2017).

Furthermore, the IoT can also reshape services in other industries, such as the insurance ecosystem, e.g., by providing insurance pay-as-you-drive (Desyllas and Sako 2013).

We also consider this cluster as intertwined, since it reflects the current situation in many ecosystems, where innovative service providers try to break up and recombine the value creation of traditional manufactures, which in turn forces them to stay innovative. This behavior can also be observed in the financial or insurance industry (Puschmann 2017).

The fourth cluster, **Data Prediction and Monitoring**, contains organizations from two ecosystems, the insurance and IIoT ecosystem. Using the LSA Algorithm, we identify seven topics from which we extract the first three keywords: 1) software, develop, manage, 2) data, software, analytics, 3) manage, claim, AI, 4) business, process, claim, 5) claim, develop, analytics, 6) manage, base, property, 7) agent, AI, global.

Both of the organizations' assigned roles in this cluster refer to using data for monitoring or predicting purposes. This is also supported by the keywords of the identified topics.

Topic 3 and 7 particularly suggest that some of the organizations use AI for data science. Hence, the organizations connected to these topics use machine learning for price optimization, e.g., for the calculation of an adequate price regarding a specific risk in the case of insurances (Desyllas and Sako 2013). Other use cases are personalized marketing, e.g., based on the lifestyle or social media activities of potential customers, customer segmentation, e.g., according to their financial sophistication, age, location, and attitude or risk assessment (Desyllas and Sako 2013; Dorfleitner et al. 2017).

Moreover, topic 4 shows that some organizations in this cluster monitor and predict the outcome of business processes. For example, insurance companies are increasingly interested in claims predictions to calculate potential financial losses. Further examples are organizations like Cognotect and Fraugster.

Figure 3 shows Cluster 04 as central and connected to almost all other clusters. Additionally, Cluster 04 is assigned to a service-oriented and product-oriented ecosystem, which makes it a potential candidate for a future core cluster, such as Cluster 01 is now.

The fifth cluster, **Broker and Agents**, again contains organizations from two ecosystems, the insurance and financial industry with two generic roles: Robo Advisors / Digital Brokers and Personal Financial Management. Using the LSA algorithm, we can define three topics from which we extract the first three keywords: 1) business, online, management, 2) business, agency, way, 3) brokerage, business, automatic.

The topics suggest that the organizations in the cluster refer to agents or brokers that are either automating and/or

managing the business. This cluster shows a typical transition from offline to online services due to digital technologies (see Kaltenecker et al. 2015).

Figure 3 shows that Cluster 05 is strongly connected to Cluster 02 “Digital Financial Services” and Cluster 04 “Data Prediction and Monitoring”. This, together with the first keywords, suggest that the included organizations offer digital services that either manage or automate financial services of organizations or investment banking. Examples in this cluster are organizations like FondsFinanz and Euroassekuranz.

Ecosystem-specific clusters

Ten of the identified clusters solely contain organizations from one ecosystem. Therefore, we consider these clusters as ecosystem-specific. We identified three automotive clusters (Cluster 06, 10, and 11), three financial clusters (Cluster 07, 08, and 14), two insurance clusters (Cluster 09, and 15), and two blockchain clusters (Cluster 12 and 13), which are discussed in this order subsequently.

Cluster 06 consists of the five topics marketplace, parts, peer, sale/dealership/operations, and base/buy/use. It contains automotive organizations connected to the generic roles Car Dealer, Mobility Service Aggregator, Parts Provider, Car Service Provider, OEM, Intelligent Infrastructure Provider, Mobility Service Platform, Car Rental Provider, and Public Transportation Provider. These roles represent the core roles of the traditional automotive industry and its components of individual transportation, such as by (rental) car, and public transportation. Notably, some of the traditional organizations are already developing digital services or platforms, such as myTaxi, the taxi hail service of Mercedes Benz. However, these services are not industry-independent as the cloud-based services found in Cluster 01.

Cluster 10 consists of topics connected to supplier/global/component, or China. It contains automotive organizations from the generic roles Tier 1–3 Supplier, and Car Service Provider. Therefore, this cluster contains organizations that produce and distribute necessary components for the car manufacturers (or OEMs). Also, these organizations seem to have a unique value creation, as the clustering did not detect similarities to organizations in other ecosystems.

Cluster 11 consists of the topics manage/base/administration, and outsourcing. It contains only automotive organizations from the generic role Value Added Service Provider. These organizations build digital services specific to the automotive industry, which can be accessed before, during, or after transportation. Examples are telematics services, navigation software, or intelligent driving assistance software.

Cluster 07 consists of the topics investment/online/social. It contains organizations from the financial ecosystem with the generic roles Stock Market, Robo-Advice/Portfolio Management, Crowdfunding, and Social Trading. Hence,

with the stock market and portfolio management, the organizations in this cluster are core components of the financial ecosystem. However, also emerging actors that base their services on innovative technologies, such as social trading and crowdfunding can be found here. On common ground, these organizations enable access to capital or the capital market.

Cluster 08 consists of similar topics and also contains organizations only from the financial ecosystem. There, the generic roles Loans, Crowdlending, and Money Transfer can be found here. The generic role Money Transfer consists of organizations such as Western Union, which offers a global network of cash transfer independent from other banks. The rest of the organizations in this cluster provide money to customers in two ways: traditional banks offer loans, and emerging players offer crowd lending platforms to connect creditors with debtors.

Cluster 14 consists only of organizations with the generic role Alternative Payment Solution originated in the financial ecosystem. The role is an umbrella for Fintechs whose applications and services concern payment and mobile payment transactions, such as the Sweden-based startup Klarna.

Cluster 09 consists of the topic compare/comparison/online. It contains only organizations from the insurance ecosystem with the generic roles Comparison Platform and Cross-Sellers. Comparison platforms, such as Getinsured or Impacthealth enable customers to form adequate decisions regarding different products and providers. Cross-seller, such as Simpleurance target the potential of selling further insurance to customers in digital environments, such as online shops, for example through selling luggage insurances after buying a flight ticket.

Cluster 15 consist of only one generic role from the insurance ecosystem, digital business services. These are services handled by external service providers in many aspects of the insurance ecosystem, including consulting, human resource management, and debt collection services.

Cluster 12 consists only of organizations connected to the generic role Blockchain Application Provider, which provide services for various topics, including invoice, e-commerce, e-payment. These organizations offer applications that are linked to on-chain services, such as voting, tokenization, asset linkages and naming registrations provided by decentralized applications. Moreover, some also offer off-chain services, where value is moved outside the blockchain.

Ultimately, cluster 13 consists of organizations in the generic roles Blockchain Platform Providers, underlying the applications of the prior cluster. Hence, these organizations build the technological basis to build, run, and test applications and thus extend the functionality of infrastructure elements. This also includes smart contract languages and scripts, testing tools, sandbox environments, integrated development environments, and frameworks for software development.

Discussion

Theoretical contribution

Based on this study, three theoretical contributions have come to light.

First, we propose a methodological innovation to analyze ecosystems, by combining the strengths of conceptual modeling and cluster analysis. Thereby, we apply and expand the work of Basole et al. (2018), and use text mining to identify core, intertwined, and ecosystem-specific clusters in ecosystems based on the similarities of the corresponding organizations, after we modeled the ecosystems consisting of generic roles and relationships between them following the approach of Riasanow et al. (2017). Next, we mapped the organizations in the 15 clusters to the previously identified generic roles. This aggregation allows to abstract from single organizations, and discuss and analyze ecosystems according to the value proposition of generic roles, which is particularly helpful for large ecosystems.

Second, in performing the clustering, we address the call for examining ecosystems as a structural entity (Adner 2017), particularly through identifying core and intertwined clusters. Cluster 01 provides three core roles, found in each of our ecosystems: cloud application, platform, and infrastructure providers, which offer industry-independent services and are, thus, the foundation of many platform ecosystems. These roles also highlight the transformation from on premise to cloud services as a central aspect of digital transformation in platform ecosystems. Thereby, we extend prior work of Kaltenecker et al. (2015), who analyzed five case studies in this context. The availability of cloud infrastructure eventually led to a plethora of digital services in all of our examined ecosystems. Notably, this cluster also identifies roles connected to data protection, digital identity, hardware providers, and regulatory authorities, as core of platform ecosystems. Clusters 02–05 show intertwined clusters, with roles from more than one ecosystem. These roles have the potential to become a core role in the future, as they demonstrated to be relevant in more than one ecosystem. For example, Cluster 02 and 05 show the critical role of digital financial services for platform ecosystems, e.g., also in the automotive ecosystems. As an example, Uber exclusively manages the payment of rides via digital financial services, such as PayPal, so that the customer does not have to bring cash. Cluster 03 shows that IIoT solution providers are breaking up and recombining the value creation of automotive OEMs (see Weill and Woerner 2015). Thereby, they are increasing the number of different services in an ecosystem and forcing the established players to innovate in order to stay competitive. Cluster 04 identified machine learning as one of the new technologies driving innovation in platform ecosystems. In our case, we observed this phenomenon particularly in the financial and

insurance ecosystem, e.g., by enabling price optimization, personalized marketing, customer segmentation, and risk assessment. Hence, our clusters help scholars to also analyze digital transformation in other than our studied platform ecosystems. Also, the roles in the core or intertwined clusters can serve as a starting point when designing a new, or transforming an established ecosystem.

Third, the clusters reveal the underlying interconnectivity and complexity of platform ecosystems while providing important triangulated insights into ecosystem-specific differences. Moreover, as only one of the 15 clusters contains roles that can be found in all of the ecosystems, the distribution of our findings are similar to the oligopoly market structure of cloud platform providers, for example. There only a few large organizations, such as AWS or Microsoft remain to provide the hyper-scaling commodity services to a large, and growing, customer base. Or, on the other hand, most of the organizations in platform ecosystems offer specialized, complementary services, which are building on platforms. However, based on the intertwined clusters, we see that innovative roles, such as IIoT solution providers, extend their range from IIoT to the automotive ecosystem, for example, which can be understood as first step towards a commodity service.

Practical contribution

Our analyses help practitioners in four important ways.

First, decision-makers, e.g., from traditional organizations may apply our ecosystems to identify potential threats to their current market position, potential opportunities to adapt to trends, and shifts in customer needs.

Second, when designing new, or transforming established platform ecosystems, we advise practitioners to definitely account for the generic roles identified in our “core” Cluster 01, consisting of cloud and on premise infrastructure providers, and cybersecurity providers. Moreover, we also suggest analyzing the impact of the roles in the intertwined clusters (Cluster 02–05), as they represent roles that are extending their services to other ecosystems. Hence, the intertwined roles may have an effect on the particular ecosystem under investigation in the future, too.

Third, the cluster analysis is helpful to analyze the size and centrality of specific roles in platform ecosystems. This step may be helpful when analyzing the importance of specific roles for the ecosystem. For example, we found that Cluster 04 that concerns data prediction is very central to most other roles of the ecosystem. Such conclusions could not be drawn by simply using the e3 value method, or ecosystem-as-a-structure to model an ecosystem.

Fourth, our proposed method enables decision-makers to understand and analyze ecosystems from two perspectives, so that decision-makers can understand the value streams between the organizations using the e3 value

model and examine the similarity to roles of other ecosystems using the cluster analysis.

Limitations and future research

Our study is subject to limitations and raises avenues for future research.

First, even though the coding of the organizations to the generic roles follows the structured approach proposed by Mayring (2010) it is contingent on subjective coding due to the human coding process. Hence, we used a second coder and measured inter-coder reliability which was acceptable above 0.8 for each of the respective ecosystems. Further, we conducted 31 interviews with external industry experts to discuss and validate our developed ecosystems.

Second, our ecosystems are limited by the information provided by the Crunchbase database and our coding of roles. Future research could enhance our suggested method, e.g., by using web scrapping algorithms to include more information.

Third, our data-driven approach using firm-level Crunchbase data limits the clustering. For future research, it would be beneficial to include ecosystem-level data, such as API calls like Benzell, Hersh, van Alstyne, and LaGarda (2019), who investigate how firms use APIs (to incorporate external data sources in the ecosystem) and the consequences of using them on firm performance. The inclusion of more ecosystem-level data may shed more light on the underlying dynamics of an ecosystem. By using APIs, future research could examine technical, instead of economic, interactions in an ecosystem, for example, connected to data or knowledge sharing on the ecosystem level. Future research with ecosystem-level data can be beneficial to highlight centralities of the individual roles or organizations in an ecosystem, e.g., to identify gateways or bottlenecks.

Fourth, our data-driven approach limits the interpretability of the identified clusters. Contrary to our chosen data-driven approach, future research takes a theory-driven stance to verify the clusters based on the perspective of business ecosystems, business networks, and inter-firm relationships literature. We hope this reveals additional, or complementary insights to our findings. Therefore, we suggest developing a framework or taxonomy for clusters in platform ecosystems that integrate the respective literature streams, such as the role of boundary resources, the balance between value creation and value capture, or various different relationships of actors in an ecosystem.

Ultimately, our analysis is limited to platform ecosystems. We suggest future research to apply our method in the context of other types of business ecosystems.

Conclusion

This work provides a new method to analyze ecosystems, and identify similarities of the roles in the ecosystems, by combining the strengths of conceptual modeling using e3 value and cluster analysis using text mining. Applying this method to the automotive, blockchain, financial, insurance, and IIoT ecosystems, we found 15 organizational clusters based on the similarity of the containing organizations. Among these clusters, we identified one core cluster that contains organizations occurring in all ecosystems, and four intertwined clusters containing organizations from at least two ecosystems. Our work contributes to ecosystem theory and the phenomenon of digital transformation on an ecosystem level in multiple ways. We encourage scholars, platform owners, and complementors that seek to analyze platform ecosystems to adapt our approach.

Acknowledgements The authors would like to thank all anonymous reviewers and the editors for their helpful comments and suggestions. This research is partly funded by the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) as part of the ‘Collaborative Research Center 768: Managing cycles in innovation processes – Integrated development of product service systems based on technical products’ (TP C1).

Funding Information Open Access funding provided by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adner, R. (2006). Match your innovation strategy to your innovation ecosystem. *Harvard Business Review*, 84(4), 98–109.
- Adner, R. (2017). Ecosystem as structure: An actionable construct for strategy. *Journal of Management*, 43(1), 39–58.
- Agarwal, R., Gao, G. G., DesRoches, C., & Jha, A. K. (2010). The digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, 21(4), 796–809.
- Altman, E. J., & Tuschman, M. L. (2017). Platforms, open / user innovation, and ecosystems: A strategic leadership perspective. Harvard Business School Organizational Behavior Working Paper No. 17-076, 177–207.
- Arreola Gonzalez, A., Pfaff, M., & Krcmar, H. (2019). Value modeling for ecosystem analysis. *Computers*, 8(3), 1–15.
- Autio, E., & Thomas, L. D. W. (2014). Innovation ecosystems: Implications for innovation management. In M. Dogson, N.

- Philips, & D. M. Gann (Eds.), *Oxford Handbook of Innovation Management* (pp. 204–228): Oxford University Press.
- Basole, R. C. (2009). Visualization of Interfirm relations in a converging Mobile ecosystem. *Journal of Information Technology*, 24(2), 144–159.
- Basole, R. C., Huhtamäki, J., Still, K., & Russell, M. G. (2016). Visual decision support for business ecosystem analysis. *Expert Systems with Applications*, 65, 271–282.
- Basole, R. C., & Karla, J. (2011). On the evolution of Mobile platform ecosystem structure and strategy. *Business & Information Systems Engineering*, 3(5), 313–322.
- Basole, R. C., Park, H., & Chao, R. O. (2018). Visual analysis of venture similarity in entrepreneurial ecosystems. *IEEE Transactions on Engineering Management*, 1–15.
- Benzell, S., Hersh, J. S., van Alstyne, M., & LaGarda, G. (2019). The paradox of openness: Exposure vs. efficiency of APIs (working paper). Retrieved at <https://ssrn.com/abstract=3432591> or <https://doi.org/10.2139/ssrn.3432591>
- Blondel, V., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 1–12.
- Bogner, A., Littig, B., & Menz, W. (2009). Introduction: Expert interviews—An introduction to a new methodological debate *Interviewing experts* (pp. 1–13): Springer.
- Böhm, M., Koleva, G., Leimeister, S., Riedl, C., & Krcmar, H. (2010). *Towards a Generic Value Network for Cloud Computing*. Paper presented at the international workshop on grid economics and business models, Ischia, IT.
- Boudreau, K. J., & Lakhani, K. R. (2009). How to manage outside innovation. *Sloan Management Review*, 50(4), 68–77.
- Bullock, R., & Tubbs, M. E. (1990). A case meta-analysis of Gainsharing plans as organization development interventions. *The Journal of Applied Behavioral Science*, 26(3), 383–404.
- Ceccagnoli, M., Forman, C., Huang, P., & Wu, D. J. (2014). Digital platforms. *Communications of the ACM*, 57(2), 38–39.
- Cusumano, M. A., & Gawer, A. (2002). The elements of platform leadership. *Sloan Management Review*, 43(3), 51–58.
- Datté, B., Alexy, O., & Autio, E. (2018). Maneuvering in poor visibility: How firms play the ecosystem game when uncertainty is high. *Academy of Management Journal*, 61(2), 466–498.
- Deerwester, S., Dumais, S., Landauer, T., Furnas, G., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society of Information Science*, 41(6), 391–407.
- Desyllas, P., & Sako, M. (2013). Profiting from business model innovation: Evidence from pay-as-you-drive auto insurance. *Research Policy*, 42(1), 101–116.
- Dorflleitner, G., Hornuf, L., Schmitt, M., & Weber, M. (2017). *FinTech in Germany*. Cham, Switzerland: Springer International Publishing AG.
- Dumais, S. T. (1991). Improving the retrieval of information from external sources. *Behavior Research Methods, Instruments, & Computers*, 23(2), 229–236.
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550.
- Eisenhardt, K. M., & Graebner, M. E. (2007). Theory building from cases: Opportunities and challenges. *Academy of Management Journal*, 50(1), 25–32.
- Evans, D. S. (2003). The antitrust economics of multi-sided platform markets. *Yale Journal on Regulation*, 20(235), 327–379.
- Fichman, R. G. (2014). Digital innovation as a fundamental and powerful concept in the information systems curriculum. *MIS Quarterly*, 38(2), 329–353.
- Fortunato, S., & Barthelemy, M. (2007). Resolution limit in community detection. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), 36–41.
- Goldberg, M., Satzger, G., & Fromm, H. (2016). Adapting IT Service Management for Successful Multi-Sourcing Service Integration. Paper presented at the proceedings of the 24th European conference on information systems (ECIS), Istanbul, Turkey.
- Gordijn, J., & Akkermans, J. (2003). Value-based requirements engineering: Exploring innovative E-commerce ideas. *Requirements Engineering*, 8(2), 114–134.
- Gordijn, J., Petit, M., & Wieringa, R. (2006). Understanding business strategies of networked value constellations using goal-and value modeling. Paper presented at the 14th IEEE international conference on requirements engineering, Beijing, China.
- Greineder, M., Riasanow, T., Böhm, M., & Krcmar, H. (2019). The generic InsurTech ecosystem and its strategic implications for the digital transformation of the insurance industry. Paper presented at the 40th GI EMISA, Tutzing, Germany.
- Hagiu, A. (2014). Strategic decisions for multisided platforms. *MIT Sloan Management Review*, 55(2), 71–80.
- Hansen, R., & Sia, S. K. (2015). Hummel's digital transformation toward Omnichannel retailing: Key lessons learned. *MIS Quarterly Executive*, 14(2), 51–66.
- Hein, A., Schrieck, M., Riasanow, T., Soto Setzke, D., Böhm, M., Wiesche, M., & Krcmar, H. (2019). Digital Platform Ecosystems. *Electronic Markets*. <https://doi.org/10.1007/s12525-019-00377-4>
- Jacobides, M. G., Cennamo, C., & Gawer, A. (2018). Towards a theory of ecosystems. *Strategic Management Journal*, 39(8), 2255–2276.
- Kaltenecker, N., Hess, T., & Huesig, S. (2015). Managing potentially disruptive innovations in software companies: Transforming from on-premises to the on-demand. *Journal of Strategic Information Systems*, 24(4), 234–250.
- Karimi, J., & Walter, Z. (2015). The role of dynamic capabilities in responding to digital disruption: A factor-based study of the newspaper industry. *Journal of Management Information Systems*, 32(1), 39–81.
- Kirshnan, V., & Ulrich, K. T. (2001). Product development decisions: A review of literature. *Management Science*, 47(1), 1–21.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.
- Lucas, H. C., & Goh, J. M. (2009). Disruptive technology: How Kodak missed the digital photography revolution. *Journal of Strategic Information Systems*, 18(1), 46–55.
- Marra, A., Antonelli, P., Dell'Anna, L., & Pozzi, C. (2015). A network analysis using metadata to investigate innovation in clean-tech—implications for energy policy. *Energy Policy*, 86, 17–26.
- Mayring, P. (2010). *Qualitative Inhaltsanalyse*. Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.
- Metzler, D., Dumais, S., & Meek, C. (2007). Similarity measures for short segments of texts. Paper presented at the European Conference on Information Retrieval.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Newbury Park, CA: Sage.
- Moore, J. (1993). Predators and Prey: A New Ecology of Competition. *Harvard Business Review*, 71(3), pp. 75–86.
- Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, 17(1), 2–26.
- Newman, M. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 8577–8582.
- Piccinini, E., Hanelt, A., Gregory, R. W., & Kolbe, L. M. (2015). Transforming industrial business: The impact of digital transformation on automotive organizations. Paper presented at the 36th international conference on information systems (ICIS), Fort Worth, TX.
- Porter, M. F. J. P. (1980). An algorithm for suffix stripping. *14(3)*, 130–137.
- Puschmann, T. (2017). Fintech. *Business Information Systems Engineering (BISE)*, 59(1), 69–76.

- Ramos, J. (2003). *Using TF-IDF to determine Word Relevance in Document Queries*. Paper presented at the proceedings of the first instructional conference on machine learning, Piscataway, NJ.
- Remané, G., Hanelt, A., Wiesboeck, F., & Kolbe, L. M. (2017). Digital maturity in traditional industries: An exploratory analysis. Paper presented at the 25th European conference on information systems (ECIS), Guimarães, Portugal.
- Riasanow, T., Burckhardt, F., Soto Setzke, D., Böhm, M., & Krcmar, H. (2018a). The generic Blockchain ecosystem and its strategic implications. Paper presented at the 24th Americas conference on information systems (AMCIS), New Orleans, LA.
- Riasanow, T., Flötgen, R. J., Soto Setzke, D., Böhm, M., & Krcmar, H. (2018b). The generic ecosystem and innovation patterns of the digital transformation in the financial industry. Paper presented at the 22nd Pacific Asian conference on information systems (PACIS), Yokohama, Japan.
- Riasanow, T., Galic, G., & Böhm, M. (2017). Digital transformation in the automotive industry: Towards a generic value network. Paper presented at the 25th European conference on information systems (ECIS), Guimarães, Portugal.
- Riasanow, T., Soto Setzke, D., Böhm, M., & Krcmar, H. (2019). Clarifying the notion of digital transformation: A Transdisciplinary review of literature. *Journal of Competences, Strategy and Management*, 10, 5–31.
- Rochet, J.-C., & Tirole, J. (2003). Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4), 990–1029.
- Röder, M., Roth, A., & Hinneburg, A. (2015). *Exploring the Space of Topic Coherence Measures*. Paper presented at the proceedings of the eighth ACM international conference on web search and data mining, New York, NY.
- Sarker, S., Sarker, S., Sahaym, A., & Bjorn-Andersen, N. (2012). Exploring value Cocreation in relationships between an ERP vendor and its partners: A revelatory case study. *MIS Quarterly*, 36(1), 317–338.
- Simpson, T., Maier, J., & Mistree, F. (2001). Product platform design: Method and application. *Research in Engineering Design*, 13(1), 2–22.
- Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: How the technology behind Bitcoin is changing money, business and the world*. New York, NY: Penguin Random House.
- Teece, D. J. (2007) Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal* 28(13), 1319–1350.
- Thomas, L. D. W., Autio, E., & Gann, D. M. (2014). Architectural leverage: Putting platforms in context. *Academy of Management Perspectives*, 28(2), 198–219.
- Tiwana, A. (2014). *Platform ecosystems: Aligning architecture, governance, and strategy*. Burlington, MA: Morgan Kaufmann.
- Tiwana, A. (2015). Evolutionary competition in platform ecosystems. *Information Systems Research*, 26(2), 266–281.
- 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.
- Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research*, 37, 141–188.
- Urmetzer, F., Gill, A., & Reed, N. (2018). Using business ecosystems mapping to generate new competitive value propositions. Paper presented at the CIE 48, Auckland, NZ.
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *Journal of Strategic Information Systems*, 28(2), 118–144.
- Weber, M. L., & Hine, M. J. (2015) Who Inhabits a Business Ecosystem? The Technospecies as a Unifying Concept. *Technology Innovation Management Review* 5(5), 31–44.
- Weill, P., & Woerner, S. L. (2015). Thriving in an increasingly digital ecosystem. *MIT Sloan Management Review*, 56(4), 27–34.
- Westerman, G., Bonnet, D., & McAfee, A. (2014). *Leading digital: Turning technology into business transformation*. Boston, MA: Harvard Business Review.
- Westerman, G., Calmèjane, C., Bonnet, D., Ferraris, P., & McAfee, A. (2011). *Digital Transformation: A Road-Map for Billion-Dollar Organizations*. Retrieved from
- Yin, R. K. (2014). *Case study research* (5th ed.). Thousand Oaks, CA: SAGE Publications Inc..
- Yoo, Y., Boland Jr., R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, 23(5), 1398–1408.
- Yu, Y., & Perotti, V. (2015). Startup tribes: Social network ties that support success in new firms. Paper presented at the proceedings of 21st Americas conference on information systems (AMCIS 2015), Puerto Rico.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.