



Systematizing the lexicon of platforms in information systems: a data-driven study

Christian Bartelheimer¹ · Philipp zur Heiden¹ · Hedda Lüttenberg² · Daniel Beverungen¹

Received: 22 January 2021 / Accepted: 3 February 2022 / Published online: 30 March 2022
© The Author(s) 2022

Abstract

While the Information Systems (IS) discipline has researched digital platforms extensively, the body of knowledge appertaining to platforms still appears fragmented and lacking conceptual consistency. Based on automated text mining and unsupervised machine learning, we collect, analyze, and interpret the IS discipline's comprehensive research on platforms—comprising 11,049 papers spanning 44 years of research activity. From a cluster analysis concerning platform concepts' semantically most similar words, we identify six research streams on platforms, each with their own platform terms. Based on interpreting the identified concepts vis-à-vis the extant research and considering a temporal perspective on the concepts' application, we present a lexicon of platform concepts, to guide further research on platforms in the IS discipline. Researchers and managers can build on our results to position their work appropriately, applying a specific theoretical perspective on platforms in isolation or combining multiple perspectives to study platform phenomena at a more abstract level.

Keywords Platform · Text mining · Machine learning · Data communications · Interpretive research · Systems design and implementation

JEL classification L86

Introduction

Over nearly 50 years, the Internet has enabled the emergence of an ever-increasing variety of platforms, while the Internet itself evolved from a digital information platform to a digital communications platform. Digital platforms are defined as a set of digital resources enabling interactions between actors while creating value (Boudreau & Hagiu, 2009; Constantinides et al., 2018; Parker et al., 2017). They manifest as boundary objects (Star, 2010; Star & Griesemer, 1989) that enable organizations and individuals to network with each

other, and through which business models and markets can be fundamentally disrupted (Iansiti & Lakhani, 2014).

Information Systems (IS) research takes two major perspectives on phenomena related to platforms—especially digital platforms—a technical/engineering perspective and an economic perspective (Gawer, 2014). From a technical perspective, platforms are built on an “extensible codebase of a software-based system that provides core functionality shared by the modules that interoperate with it and the interfaces through which they interoperate” (Tiwana et al., 2010, p. 676). In this perspective, IS scholars investigate the design (e.g., Spagnoletti et al., 2015), (third-party) development (e.g., Ghazawneh & Henfridsson, 2013), and architecture (e.g., Baldwin & Woodard, 2009) of platforms. Adopting an economic perspective, researchers investigate the economic effects and mechanisms of value creation on two- or multi-sided markets (McIntyre & Srinivasan, 2017), define platform launch strategies (Kazan & Damsgaard, 2016; Stummer et al., 2018), and investigate the evolution of platform ecosystems (Asadullah et al., 2018; Ozer & Anderson, 2015; Tiwana et al., 2010) or of online communities (Butler et al., 2014).

Responsible Editor: Yun Wan

✉ Christian Bartelheimer
christian.bartelheimer@upb.de

¹ Paderborn University, Department of Information Systems, Warburger Str. 100, 33098 Paderborn, Germany

² Hamm-Lippstadt University of Applied Science, Production and Quality Management, Dr.-Arnold-Hueck-Straße 3, 59557 Lippstadt, Germany

The platform terms referred to in different research streams are often neither identical nor even compatible. On the one hand, IS scholars use different terms to refer to the same concept (synonyms: e.g., internet platform and online platform), while, on the other hand, they use the same terms to refer to different concepts (homonyms: e.g., service platform in the domain of service science refers to an environment enabling actors to interact and thereby co-create value, while in computer science it refers to a platform providing IT services). Due to this lexical fragmentation, researchers and practitioners struggle to consistently identify and define the characteristics of platforms and specific platform types (Sørensen et al., 2015). We propose using the metaphor of a *rhizome* to describe the current configuration of platform terms. A *rhizome*—a philosophical concept established by Deleuze and Guattari (1979)—resembles a botanic rhizome and describes a system in which different nodes are connected by various edges to form a network without hierarchies—opposing tree structures that use hierarchies.

In a rhizome, every node can and must be connected to every other node, establishing a multiplicity of each node, and allowing investigators to browse through a rhizome using a multitude of entry and exit points (Deleuze & Guattari, 1979). A rhizome can break at any point and still keep growing (Deleuze & Guattari, 1979). In IS, a rhizome has been used sporadically as a metaphor for diverse concepts, including networks (Atkinson & Brooks, 2005; Gachet & Brézillon, 2005), decision processes (Humphreys, 2021; Nolas, 2008), assemblage (Leclercq-Vandelannoitte et al., 2014; Sesay et al., 2016), discourse (Iivari et al., 2017), and transformation (Márton, 2021). We use the rhizome metaphor to describe the interplay of different research approaches on digital platforms as a complex phenomenon that is subject to lexical fragmentation and lacking distinguishable hierarchies. Still, different subtypes of digital platforms might have similar characteristics, constituting a stream of literature that yields a multitude of entry and exit points.

Even though earlier attempts have been made to structure research on the rhizomatic nature of platforms, they had three crucial and recurring shortcomings that have prevented establishing a holistic perspective. First, most related research adopted a reductionist perspective—emphasizing, for example, a technical over an economic perspective—while only a few papers took an integrative approach (e.g., Asadullah et al., 2018; Gawer, 2009; Spagnoletti et al., 2015). However, a rhizome cannot be reduced to its parts, as it represents the multiplicity of the whole and the interconnectedness of its parts (Deleuze & Guattari, 1979). Second, previous research applied specific theoretical lenses to study platforms, e.g., the sharing economy (Sutherland & Jarrahi, 2018) or online communities (Spagnoletti et al., 2015) but without integrating their respective knowledge. Concerning

the rhizomatic nature of digital platforms, current research focuses on parts of the rhizome, which again conflicts with the multiplicity and interconnectedness of all nodes. Third, previous research focused on reviewing specific types of platforms (e.g., IoT platforms, see Hein et al., 2018, or multi-sided platforms, see Boudreau and Hagiu, 2009), while not systematizing a consistent lexicon of different types of platforms. Current attempts that systematize IS research on platforms favor a particular research stream's body of knowledge while neglecting other streams, which means that the work performed by others remains unacknowledged (vom Brocke et al., 2015).

We take up the call from de Reuver et al. (2018) for “developing a typology expressing the variety of digital platforms” (de Reuver et al., 2018, p. 133). We extend this call for research to all platform terms investigated in IS. It has been argued that developing a consistent lexicon (Habermas, 2014)—constituting the structure and definitions of platform terms in IS—constitutes an important step towards consolidating concepts and theory in any field of scientific inquiry (Berente et al., 2019), including platform research in IS. Thus, we attempt to disentangle the platform rhizome, based on reducing its complexity through decomposing the IS platform literature, to identify a structure that is modular, yet connected. Modularity refers to breaking down a system into discrete pieces (i.e., modules) that communicate with each other only through standardized interfaces (Langlois, 2002), thereby decomposing a system into fine-grained, interacting subsystems that can themselves be subject to decomposition. Simon (1962) introduced the notion of (near) decomposability and viewed hierarchy as a prominent organizing principle of nature. Building on this point of view, we posit that decomposition can be a helpful strategy to analyze rhizomatic structures, too. To assemble our home discipline's knowledge (Tarafdar & Davison, 2018) and provide a forward-looking lexicon of platform research, we formulate our research question as follows: *What concepts appertaining to platforms are investigated in the IS discipline, and how do these concepts relate to each other to constitute a decomposed, forward-looking lexicon of platforms for IS research?*

We apply an inductive and data-driven research approach, inspired by other papers that restructured a different research field (e.g., Antons & Breidbach, 2018; Sakata et al., 2013), to collect, analyze, and interpret the current lexicon on platforms in IS research. In a mission to develop “computationally intensive theory” (Berente et al., 2019, p. 51), we adapt the research process outlined by Müller et al. (2016). Hence, we first collect 11,049 peer-reviewed papers—representing 95%¹—from leading IS journals to conference proceedings.

¹ As explained below, the 95% were those papers we were able to access of all papers on platforms ever published in considered IS outlets.

Second, we apply unsupervised machine learning methods to analyze these vast amounts of unstructured data and identify the 26 most influential platform terms along with each term's semantically most similar words (SSW). We then cluster the identified platform terms hierarchically, visualizing the clusters with a dendrogram. Third, we discuss the results to interpret and systematize the implications of our data-driven findings. Finally, we consolidate our findings, presenting a decomposed, forward-looking lexicon of platform concepts that can be used individually or can be combined to analyze digital platforms in IS.

From a theoretical perspective, this paper contributes unique data-driven findings, elucidating the lexicon of terms that constitutes platform research in IS. Our study is the first to collect and analyze our discipline's entire body of knowledge on platforms, covering more than 44 years of academic inquiry, dating from 1975 to 2019. Based on an inductive, data-driven approach, we identify, quantify, and systematize the most influential platform terms. Thereby, we complement non-empirical approaches that have structured the field conceptually (e.g., Fu et al., 2018; Schreieck et al., 2016; Tiwana et al., 2010) with data-driven insights resulting in a decomposed model and a lexicon of platform terms. Our results enable other researchers to comprehend the rhizomatic nature of digital platforms by using the extended lexicon of research on platforms and to position their insights vis-à-vis previous theory. Our findings provide a consistent conceptual frame of reference for research on digital platforms that systematizes and consolidates related research focusing on particular concepts and theories in this area. From a managerial perspective, our results outline how platform technologies interplay with economic effects and online communities, enabling practitioners to better understand and design strategies for value creation with platforms.

The remainder of the paper unfolds as follows. In Section 2, we justify and describe our research method. In Section 3, we report our data and their analysis in detail to make our approach transparent and replicable. In Section 4, we code, interpret, and discuss our findings against the backdrop of related research. In Section 5, we propose a decomposition of platform terms and a lexicon to systematize our discipline's rhizomatic research on platforms. As customary in many inductive studies, the presentation and discussion of related work follow the reporting of our data (Müller et al., 2016). Section 6 concludes the paper and outlines prospects for future research.

Research method

Performing a data-driven study enables us to analyze the entire body of knowledge on platforms in IS. In doing so, we build our research endeavor upon “the idea that research can

start with data or data-driven discoveries, rather than with theory” (Müller et al., 2016, p. 291). Several researchers already applied this strategy in service science (Antons & Breidbach, 2018), medicine (Churilov et al., 2005), financial analysis (Sung et al., 1999), and even fishery (Syed & Weber, 2018), to name just a few. We instantiate our data-driven approach to acquire interpretable results, instead of reaching the highest accuracy possible (Müller et al., 2016). After analyzing the data, we interpret our findings in a theory-driven discussion. Our research, therefore, covers all three phases of data-driven studies: Data collection, data analysis, and result interpretation (Müller et al., 2016).

Several methods can be used to identify platform concepts and, thereby, to systematize the lexicon of platforms, from analyzing literature, including citation and co-citation analysis (Osareh, 1996), to manual literature reviews (Webster & Watson, 2002), or text mining approaches. We opted for the latter in the light of the following advantages it entails for our endeavor. Text mining provides several methods and techniques to generate interpretable semantic representations from textual data (Miner et al., 2012). First, as a sub-field of text mining, Natural Language Processing (NLP) provides methods for automatically processing vast amounts of unstructured textual data (Miner et al., 2012), enabling the extraction of new knowledge from the data, while being scalable and reliable (W. Fan et al., 2006; Frawley et al., 1992). Its ability to analyze the entire body of knowledge published on platforms enables us to systematize the lexicon of platform research, covering more literature than researchers can read and analyze manually in a reasonable period of time (Debortoli et al., 2016). Second, NLP does not require making a-priori decisions on including or excluding papers from the analysis, eliminating the selection bias that is inherent in manual literature reviews (Indulska et al., 2012).

Figure 1 displays the six steps of data collection (step 1) and data analysis (steps 2–6). When compiling our data set (step 1), we had to make three decisions. First, as the lexicon of platforms differs for each scientific domain (e.g., platforms in engineering; Simpson, 2004), we decided to emphasize conceptual clarity over completeness, by solely analyzing platform research in IS. Second, consistent with our first decision, we again emphasized conceptual clarity over generating a bigger data set, focusing on peer-reviewed and high-quality papers. Therefore, the search process was limited to journals and conferences that had been ranked B or higher in the VHB JOURQUAL3 ranking, sub-discipline business information systems (Henning-Thurau et al., 2004), the most prominent ranking of scientific outlets for IS research in the German-speaking community. Across all journals and conference proceedings included in the analysis, the number of papers that contain the term ‘platform’ at least once reveals that the concept has been used broadly in our discipline (Table 1).

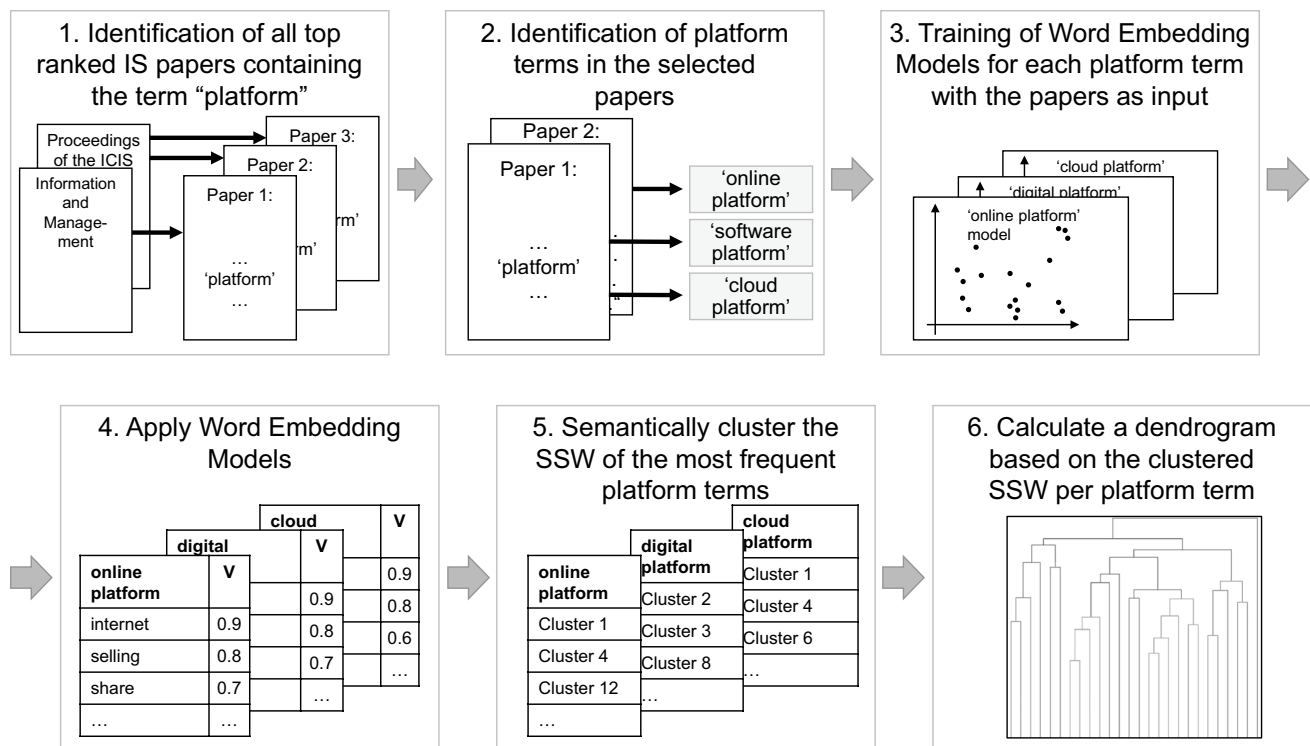


Fig. 1 Computational steps for identifying and quantifying platform terms in IS

Table 1 Distribution of the term 'platform' across selected IS journals and conference proceedings

Journal	#	Journal	#
Proceedings of the ICIS	2052	Information Systems (IS)	164
Proceedings of the ECIS	1675	Journal of Decision Systems	154
Decision Support Systems (DSS)	857	Decision Sciences	149
Information Systems Frontiers (ISF)	518	Data & Knowledge Engineering	141
Information & Management	495	Group Decision and Negotiation	120
Management Information Systems Quarterly (MISQ)	422	Communications of the ACM (CACM)	109
Proceedings of the ER	392	INFORMS Journal on Computing (JOC)	90
Information Systems Research (ISR)	374	ACM SIGMIS Database	60
Journal of Management Information Systems (JMIS)	368	ACM Computing Surveys	59
Computers and Operations Research	330	Mathematical Programming	57
Journal of Information Technology (JIT)	317	IEEE Transactions on Software Engineering	56
European Journal of Information Systems (EJIS)	272	IEEE Transactions on Engineering Management	33
Business & Information Systems Engineering (BISE)	249	ACM Transactions on Information Systems	21
International Journal of Electronic Commerce (IJEC)	241	ACM Transactions on Computer-Human & Interaction	20
The Journal of Strategic Information Systems (JSIS)	226	Information and Organization	15
Information Systems Journal (ISJ)	213	SIAM Journal on Computing	14
Artificial Intelligence	207	ACM Transactions on Management Information Systems	13
MIS Quarterly Executive	207	Journal of the ACM (JACM)	3
Electronic Markets (EM)	189	Journal of Computational Finance	0
Journal of the Association for Information Systems (JAIS)	167		

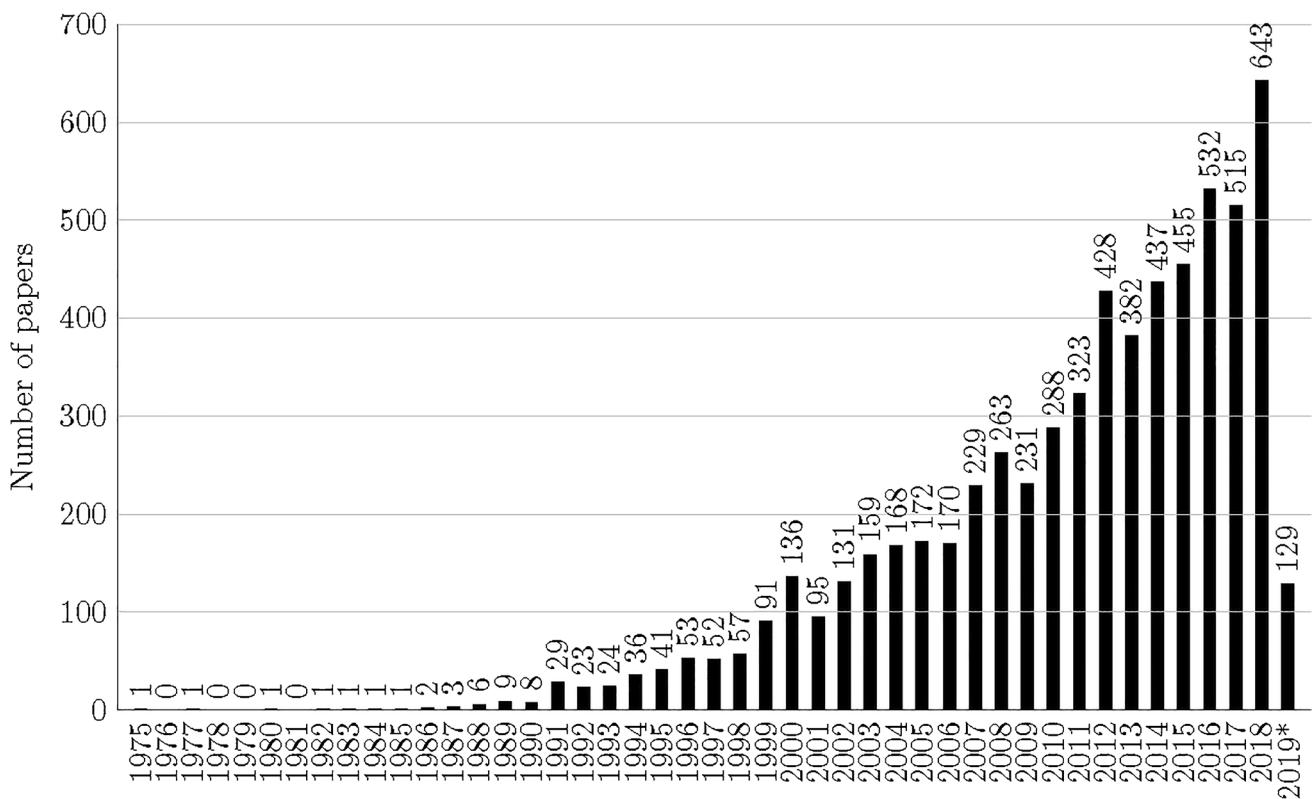


Fig. 2 Number of papers considering platforms, categorized per year

Third, we did not restrict the year range and included publications from the entire research history on platforms in IS in order to present a full picture of historic and recent understandings of platforms. As seen in Figure 2, platforms are a phenomenon that is of increasing interest to the IS discipline. The whole period, from 1975 to 2019, saw no historic disruption of the foundational concepts of platforms. In total, we identified 11,646 papers, of which we were able to access and analyze almost 95% (11,049 papers).

Figure 2 shows that platforms are receiving an increasing amount of attention in the IS discipline. The number of papers on platforms has roughly doubled every seven years, especially in the latter two decades.

We implement a bottom-up coding approach (Urquhart, 2013) to identify platform terms without introducing the bias usually caused by building on extant theory (step 2). Instead, we put aside theory in favor of identifying data-driven results with NLP. First, we tokenized the text (Manning & Schütze, 1999) to identify individual words from the papers contained in the data set. Second, we searched for any occurrences of the word ‘platform’ and extracted them along with the three tokens occurring before the term ‘platform’. In order to leave out terms that do not convey semantic value, we defined function words (stopwords) and

excluded them from the results (Gupta & Lehal, 2009; Munková et al., 2013; Pennebaker et al., 2003).

Trimming the results of function words and punctuation resulted in identifying the pure platform terms referred to in a paper. However, we took great care not to remove function words before the full platform terms were extracted, to avoid distorting their meaning. Third, we unified the lexical forms of the platform terms we extracted (Vijayarani et al., 2015) by lemmatization (i.e., restoring the word to its dictionary form; Manning & Schütze, 1999). We did not lemmatize the whole phrase, however, because there is no research available upon which we can build to define all lexical forms of the platform terms. For instance, it is debatable whether a ‘collaboration platform’ is the same as a ‘collaborative platform,’ both of which would have led to the same word stem with lemmatization. The second step resulted in identifying the number of papers that mention each platform term we extracted, before highlighting those terms that are used most frequently in the IS discipline.

In order to prepare our dataset for analysis, we performed additional pre-processing steps (step 3). First, we analyzed the whole text of the papers regarding the context in which each platform term occurred. We then concatenated the parts of each platform term using underscores, which allowed us to handle them as one word. Additionally, we lemmatized

and tokenized the words in an additional pre-processing step to identify the best possible results as an input for the following steps. We then used the resulting data as an input for performing machine-learning approaches to analyze latent structures and dependencies in our data set. To gain a deeper understanding of the terms, we trained and applied a Word Embedding Model (WEM) to each of the platform terms. WEMs are neural network vector representations of words that describe their semantic similarity (Mikolov et al., 2013). This approach assumes that a word's meaning is similar to the meaning of the words occurring in its immediate proximity. For our analysis, we applied the word2vec model (Goldberg & Levy, 2014) which, as an implementation of the concept proposed by Mikolov et al. (2013), has been proven to provide high-quality results (Lilleberg et al., 2015; Thomas & Azhuvath, 2018).

Based on the words occurring in their proximity, we calculated the similarity vector for each platform term (step 4), again applying the word2vec model. From this analysis, we retrieved the semantically similar words (SSW) for each platform term. The SSW enabled us to elucidate the semantic meaning of each platform term in more detail. Since the model only provided us with the similarity vectors of two words, we had to take the edited platform terms that were merged in step 3. We then ranked the words in descending order of their vectors, to identify the most similar words. Next, we excluded all function words, leaving us with up to 100 SSW per platform term. At this point, we also selected the platform terms to be analyzed further, since the next step would take all remaining SSW that were identified across all platform terms as an input. We limited our analysis to terms that occurred in at least 150 of the identified 11,049 papers.² We defined this threshold based on running the analysis several times, using different thresholds and inspecting the resulting SSW for validity and interpretability.

Since the SSW of the remaining terms might still have overlapped semantically, we performed k-means clustering for all relevant SSW (step 5)—one of the most popular, simple, and efficient algorithms used for clustering. As k-means clustering requires a numerical input, we again used a WEM to compile a semantic representation of the identified terms. This time, though, we used a pre-trained model, striving for a general semantic orientation of the words, and not for an orientation that is specific to the context of IS publications. We employed a publicly available model that has been trained on 100 billion words from Google News (Google Inc, 2013). The pre-trained model provided a numerical vector value, which we used as an input for k-means clustering, aimed at reducing the number of words to a cluster size of

10% of the words, so as to not overfit the results. The result of this step was a list of clusters for each platform term.

Completing the lexical framing, we clustered the identified platform terms hierarchically and visualized the clusters as a dendrogram (step 6). Hierarchical clustering is a two-stage process consisting of vector distance calculation and cluster calculation. In our case, each vector represented one platform term and contained the identified clusters of their SSW as dimensions. Thus, the value of each dimension equaled the number of occurrences of one cluster for a particular platform term. We then adopted average linkage (Almeida et al., 2007) to compute hierarchical clusters that can be visualized in dendrograms. A dendrogram is a graph consisting of edges and vertices where, in our case, each vertex represents a platform term or one cluster of platform terms. Our dendrogram reflects the rhizomatic nature of platform concepts but reduces the complexity inherent to a rhizome by identifying different lenses that can serve as entry points to the rhizome.

Having completed all data analysis activities, we proceeded with interpreting the results (Müller et al., 2016) in our quest to develop a more consistent lexicon of platform terms. First, we revisited authoritative definitions of these platform terms by identifying the most cited papers defining the terms. Second, we updated the definitions to sharpen their profile as analytical lenses that can be applied to research different aspects of digital platforms. Further research can build on this lexicon by either using one of these lenses to guide their research or by combining different lenses purposefully to investigate the complex interplay of different aspects in a specific platform. For instance, our lexicon provides clear terminology on which others can analyze AirBnB as a digital platform that also establishes a two-sided market, identifying both aspects and their interplay clearly.

Data analysis

Analyzing the resulting data set of 11,049 papers, we identified roughly 35,000 platform terms, with papers ranging from one to over 600 total platform occurrences (Gal-Or et al., 2018). In total, we identified more than 297 unique platform terms that occurred in at least ten different papers. The most frequently used platform terms are listed in Table 2, ranked in descending order of the number of papers in which they occur. A historical analysis showing the number of papers published per year for each platform term every year since 1975 is available as an online appendix.

During the third and fourth step, we trained a WEM for each of the platform terms listed Table 2. The SSW of the

² The same analysis with different thresholds of 50, 100, 150, and 200 words can be found in the appendix.

Table 2 Platform terms used most frequently in the IS literature (bold line shows a threshold of 150 occurrences)

Platform term	papers	Platform term	papers	Platform term	papers
social media platform	649	common platform	199	collaborative platform	102
online platform	618	sharing platform	188	open source platform	101
software platform	535	sided platform	188	web platform	96
it platform	495	internet platform	182	integration platform	92
technology platform	487	multiple platform	181	social network platform	91
digital platform	422	communication platform	176	integrated platform	90
computing platform	347	social platform	175	product platform	86
mobile platform	332	crowdfunding platform	169	computer platform	83
technological platform	315	ecommerce platform	167	android platform	82
service platform	284	social networking platform	156	microblogging platform	81
hardware platform	261	cloud platform	156	webbased platform	81
development platform	222	cross platform	132	electronic platform	76
crowdsourcing platform	220	trading platform	129	information platform	73
technical platform	214	single platform	122	community platform	62
open platform	207	collaboration platform	111	simulation platform	58

platform terms—as the output of these models—consist of (1) the word in its lemmatized form, and (2) a similarity value. The similarity value is based on the cosine similarity between both terms, where ‘1’ displays full semantic similarity, and a value close to zero indicates low to no semantic similarity. An excerpt of the SSW for the term ‘software platform’ is presented in Table 3, sorted by their similarity in descending order (the SSW for each platform term are available as an online appendix).

As described previously, we set the threshold for the minimum number of papers in which a specific platform term occurs to 150 (up to ‘cloud platform’, Table 2). The remaining 26 platform terms are further investigated in the fifth and sixth step of our research process. Finally, this step of the data analysis resulted in a dendrogram (Figure 3) which depicts the platform terms and their hierarchical clustering based on their SSW, with the scale on top indicating the distance value between each pair of vertices.

Interpretation and discussion

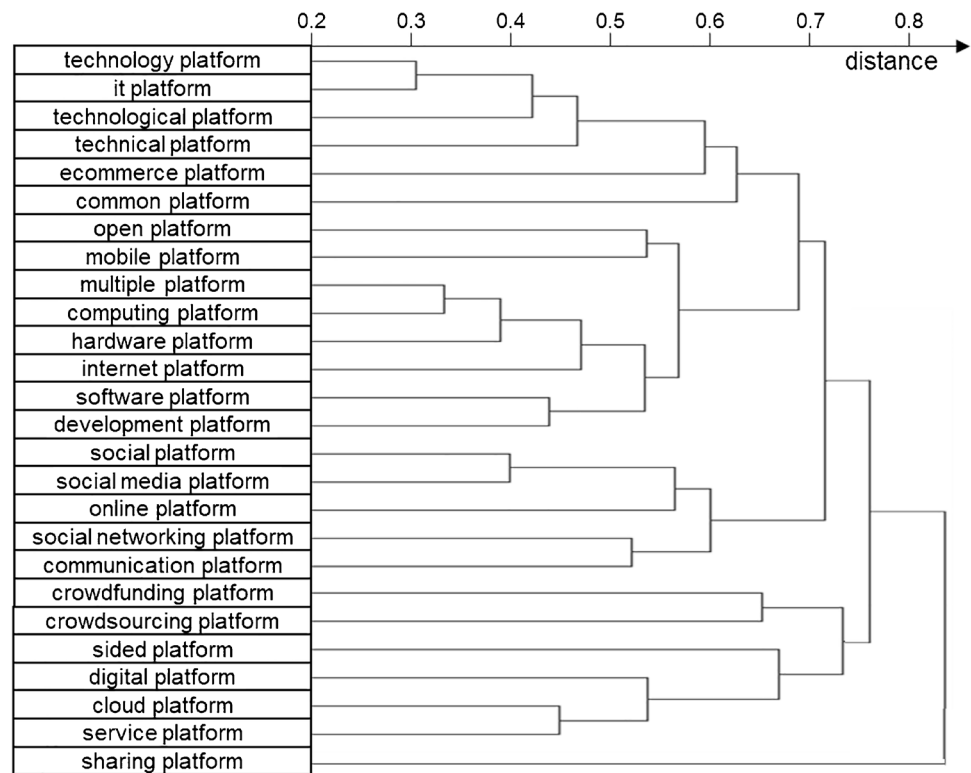
Our aim is to develop a decomposition and systematize a consistent lexicon of platform terms to guide future IS research. Building on our data, we now interpret and discuss the clustering of the identified terms (Berente et al., 2019). We started this process by interpreting terms that exhibit a distance of 0.6 to 0.7 to each other (Fig. 3), leading us to identify six clusters in total. Identifying six clusters was a normative decision. Fewer clusters would have required combining clusters that appeared to be different (e.g., ‘social platform’ and ‘computing platform’), while more clusters would have obliterated the similarities between them (e.g., ‘social platform’ and ‘social networking platform’). Our six clusters are (1) *abstract technology views on platforms*, (2) *specific views on hardware and software platforms*, (3)

social communities and online platforms, (4) *economic platforms as digital markets*, (5) *general properties of platforms as IS artifacts for value co-creation*, and (6) *sharing*

Table 3 Semantically most similar words, exemplified for ‘software platform’

SSW	similarity
platform	0,614
software	0,565
linux	0,557
api	0,546
application	0,541
mashup	0,527
dis	0,526
thirdparty	0,525
deployment	0,517
kit	0,516
iot	0,514
brokering	0,512
platforms	0,511
suite	0,508
cordova	0,508
third-party	0,507
sdk	0,506
apache	0,503
middleware	0,502
multicore	0,499
symbian	0,497
mobile	0,496
ios	0,490
toolkit	0,486
ledger	0,480
windows	0,479
vidis	0,479
ocean	0,479
complement	0,478
intranet	0,474

Fig. 3 Dendrogram, visualizing distances among the most common platform terms in IS research



platforms. In the following, we interpret and discuss the terms from each of the six clusters, focusing on the terms' position in the dendrogram (Fig. 3), their SSWs (presented in the online appendix), seminal definitions in the IS literature (presented in the online appendix), and their use over time. Figure 4 visualizes the number of papers published per year for every relevant platform term identified, structured by the six clusters (top left to bottom right).

Abstract technology views on platforms

The first cluster comprises the terms 'technology platform,' 'IT platform,' 'technological platform,' 'technical platform,' 'ecommerce platform,' and 'common platform,' all of which refer to technical aspects of platforms. 'Technology platform' is a superordinate/collective concept that describes platforms from a technical perspective on an abstract level and its utilization has been increasing in popularity among IS researchers since its inception. A 'technology platform' refers to various contexts related to IT and IS (e.g., Njenga & Brown, 2012; Purao et al., 2018) and is defined as "a set of technologies that have been developed for various applications but share a common underlying basic concept" (L.-S. Fan et al., 2015, p. 2). Thus, the term refers to "a set of design elements and interfaces that make up a technology" (Kraemer &

Dedrick, 2002, p. 9). The SSWs underline this view since they primarily consist of terms associated with the technical implementation of platforms (e.g., 'functionality,' 'application,' 'device,' 'API,' 'middleware,' 'interoperable,' 'backend,' 'modular,' 'interface'). Another term frequently used in the IS literature is 'IT platform.' Both 'technology platform' and 'IT platform' exhibit the lowest distance between all terms, since they share a substantial set of their SSWs. Unsurprisingly, 'IT platforms' are described as "the extensible codebase of a software-based system that provides core functionality shared by the applications that interoperate with it and the interfaces through which they interoperate" (Tiwana et al., 2010, p. 676). The term refers to the technical structure of platforms that features a layered modular architecture. Due to the overlap of definitions, the similarity of their meaning, and the inherent focus of the IS discipline on *information technology*—which is more specific than the term 'technology'—we propose to abandon the term 'technology platform' in favor of an 'information technology platform', to indicate that IS research concerning platforms invariably focuses on information technology.

'Technological platforms' refer to programming languages, frameworks, etc. used in Computer Science (e.g., Blechar et al., 2006; Prechelt, 2011). In IS research, the term is employed rather nonspecifically, as evidenced by its SSWs

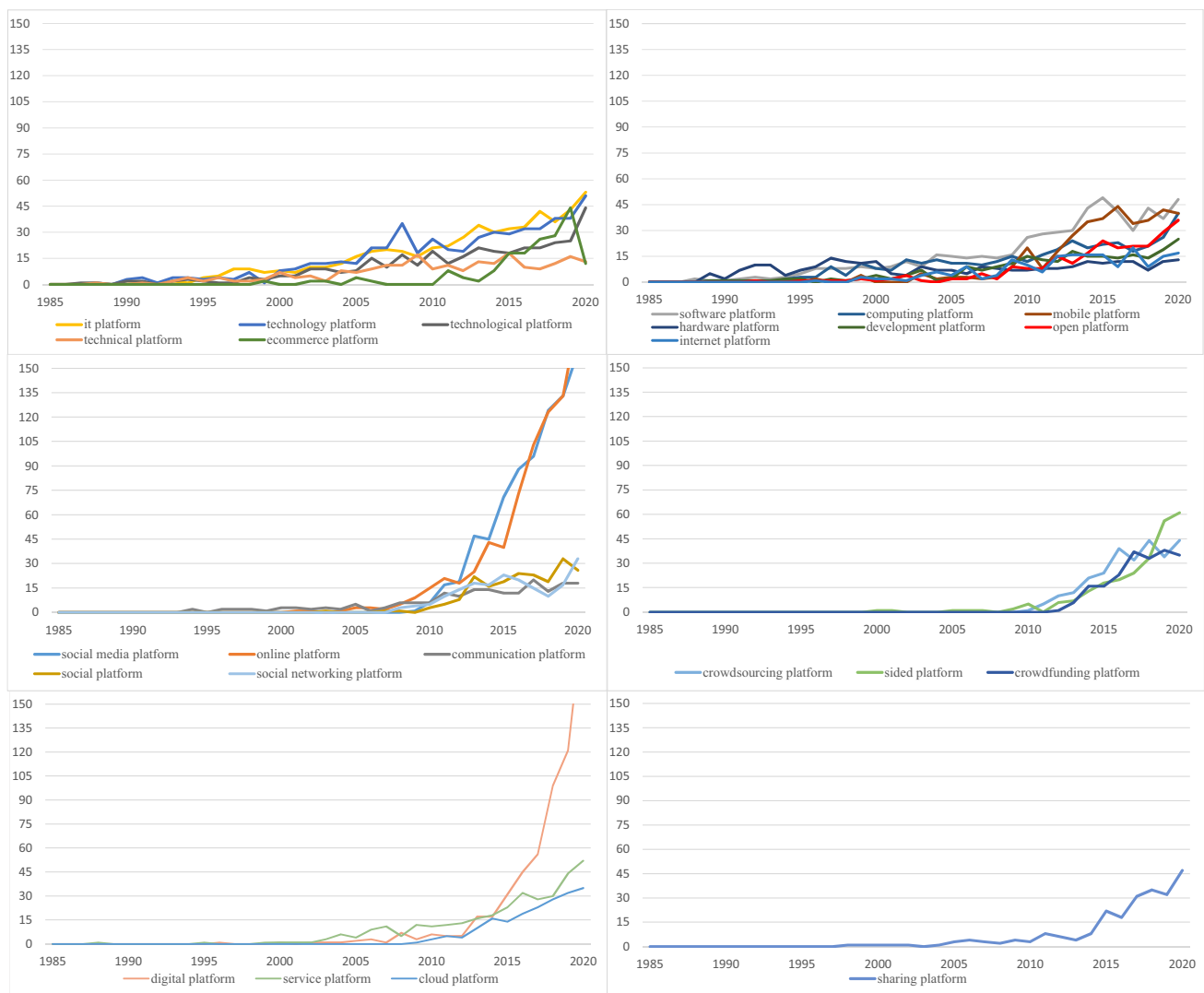


Fig. 4 Historical development of the platform terms' utilization, reported per cluster. Each diagram shows the number of papers containing a platform term published per year, for the years 1985 to 2019 (as nearly no papers were published in earlier years, cf. Figure 2)

(the ten highest-ranked SSWs are 'sophisticated,' 'option,' 'easily,' 'manner,' 'complement,' 'entire,' 'fundamentally,' 'configuration,' 'effectively,' and 'technologically'). Gawer (2014) posits:

Technological platforms can be usefully conceptualized as 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; and (3) entail a modular technological architecture composed of a core and a periphery. (p.1240)

While the term is used frequently in the recent IS literature, our triangulation of the definition by Gawer

(2014) and the SSWs reveals that the term 'technological platform' is an umbrella term that does not refer to any specific view of platforms. Hence, we recommend future researchers to discontinue its use in favor of more specific platform terms.

At the outset of platform research in IS, the term 'technical platform' was amongst the most frequently used term. However, it seems to have lost its appeal since, making it one of the least frequently used terms in this cluster (cf. Figure 4). It now seems to be confined to a technical context, in which hardware- and software-based infrastructure is described (Gillespie, 2010), primarily in the Computer Science literature (e.g., Elbanna & Linderroth, 2015; Mustonen-Ollila & Lyytinen, 2003). In this context,

Benlian et al. (2015, p. 214) define the term ‘technical platform’:

All facets of a platform related to the technical development of third-party applications including, for example, the provision of APIs and SDKs as well as all kinds of regulatory processes (e.g., quality and content checks), documentations (e.g., help files) and communications (blogs or forums in the developer community) that go along with application development. (p.214)

This definition includes all artifacts that are related to a technical view on platforms. However, because the utilization of the term is declining and its SSWs align with the SSWs of ‘technology platform’ and ‘IT platform’, we recommend not using this term. Instead, researchers could refer to the more frequently used term of ‘information technology platform’ or any other more specific platform terms.

The term ‘ecommerce platform’ (or e-commerce platform) is defined as the following:

E-commerce platform provides users with a variety of business service component [sic], through which the business service component allows users to complete the online transaction process. [...] E-commerce platforms not only provide users with functions of online transaction, but also provide users with a series of support services (Huang et al., 2011, pp. 2171–2172).

The term features the highest distance to all other terms in this cluster. Inspecting the timely distribution of the term in our data reveals that its usage seems to have peaked some years ago but is now outdated having been substituted with other terms since. Many of this term's SSWs refer to aspects that are now part of other, more specific research streams on platforms (e.g., ‘carsharing,’ ‘collaborate,’ ‘Airbnb,’ ‘crowdsourcing’). Examples for ‘ecommerce platforms’ provided in the literature (e.g., Amazon, ebay, Airbnb, Tripadvisor) substantiate this observation. Thus, we argue for discontinuing the use of the term ‘ecommerce platform’ in favor of using more specific or differentiated platform terms.

The term ‘common platform’ features a high distance to most of the other platform terms. Also, the word ‘common’ neither refers to the inner workings of a platform nor does it bear semantic value (cf. ‘most common platform’ in Park et al., 2007). Thus, we decided to exclude it from our analysis.

While this cluster contains a broad access to platforms, we view ‘IT platform’ as the most prominent term to be used in IS research when referring to a digital platform in a general sense.

Specific technology views on hardware and software platforms

The second cluster comprises the terms ‘open platform,’ ‘mobile platform,’ ‘computing platform,’ ‘hardware platform,’ ‘internet platform,’ ‘software platform,’ and ‘development platform.’ The term ‘multiple platform,’ which may have been shortened from its plural form during data preprocessing, is another term that we decided to drop due to its lack of clear semantics, as we did with ‘common platform’. Even if ‘multiple platform’ has a low distance score to ‘computing platform,’ a close inspection of the papers containing this term revealed that the term carries no specific meaning on its own. In contrast, we identified all other platform terms in this cluster as subtypes of the more general term ‘information technology platform’, as identified in cluster one.

The term ‘open platform’ does not refer to a specific platform type, but represents a research stream in IS that studies how platform *openness* impacts the development, evolution, and commercialization of a platform (Boudreau, 2010). Platform openness has been gaining increasing intention over the last couple of years, as researchers investigate it “as a governance-related concept reflecting the trade-off between retaining and relinquishing control over a platform” (Benlian et al., 2015, p. 210). The term's SSWs reveal topics of particular interest relating to, e.g., licensing, commercialization, proprietary (software), interoperability, and monetization.

The term ‘mobile platform’ is often used in the contexts of smartphones and other mobile devices (SSWs include ‘WhatsApp,’ ‘Symbian,’ ‘tablet,’ ‘smartphone’) that constitute boundary objects in mobile ecosystems that involve mobile device manufacturers, mobile network operators, mobile application developers, and other stakeholders (Basole & Karla, 2011). For a decade, the term was amongst the two most frequently used terms in this cluster. ‘Mobile platforms’ are viewed by Sørensen et al. (2015, p. 196) as:

Multi-sided markets [that] critically rely on architectural leverage (Thomas et al., 2014) through a critical mass of complementors and customers. Boundary resources (Ghazawneh & Henfridsson, 2013; Eaton et al., 2015) can support a highly distributed process subjected to combinations of centralised control and decentralised generativity (Tilson et al., 2010b).

Many papers investigate how different degrees of platform openness can lead to competitive advantage and to attracting more actors to join a platform ecosystem. Hence, “the issue of platform openness is therefore a critical issue for mobile platforms” (ibid).

The term ‘computing platform’ has been used frequently since 1995 and ranked second as the most frequently used terms in this cluster in more recent years (cf. Figure 4). The term's SSWs reveal that research on ‘computing platforms’

concerns, among others, the Internet of Things (IoT), which is the SSW with the highest similarity score. Consequently, Athanas and Abbott (1995, p. 16) have argued, back in the mid-1990s, that ‘computing platforms’ were “emerging as a class of computers that can provide near application-specific computational performance.” Other common topics identified in the SSWs are virtualization, deployment, and scalability. Hence, we conclude that the term ‘computing platform’ is used by IS scholars to report on scenarios in which computational power is outsourced to platforms that can be employed as-a-service to solve computational problems.

The term ‘hardware platform’ is used if referring to hardware issues related to platforms (SSWs include ‘installation,’ ‘mainframe,’ ‘workstation,’ ‘PC,’ ‘microcomputer’). This view is in line with definitions in the IS literature which introduce a ‘hardware platform’ as “a family of architectures that allow substantial re-use of software” (Keutzer et al., 2000, p. 1528) that “executes software application programs” (de Michell & Gupta, 1997, p. 349). In line with this definition, the term ‘software platform’ is used to refer to the development and deployment of software applications (SSWs: ‘linux,’ ‘application,’ ‘deployment,’ ‘SDK,’ ‘apache,’ ‘iOS’). The definition by Taudes et al. (2000) explains how software and hardware platforms are integrated: “A software platform is a software package that enables the realization of application systems. [...] Together with the hardware and the organizational knowledge about planning, designing, and operating application systems, the software platforms in use constitute a firm’s information technology infrastructure” (Taudes et al., 2000, pp. 227–228). However, both terms refer to separate views on platforms that need to go hand in hand to successfully design and develop platforms from a technical point of view.

Surprisingly, the IS literature does not define the next term in this cluster, ‘development platform.’ The literature only provides an example of a ‘development platform’ with the open-source development platform *Eclipse*³ (Mehra et al., 2011). This observation is in line with SSWs such as (ruby on) ‘rails,’ ‘ocean,’ ‘j2ee,’ and ‘petrel,’ all representing other programming frameworks. Thus, research using the term ‘development platform’ concerns tools and frameworks that support programmers with the software development process.

The last term in the second cluster is ‘internet platform.’ It has received relatively stable attention throughout the last decade, while all other terms, apart from ‘hardware platform,’ are used more often in the IS literature. The term refers to the prospects of the Internet connecting distinct groups of users remotely by enabling “transactions [...] by using the Internet platform (e.g. TCP/IP, HTTP, XML) in

conjunction with the existing IT infrastructure” (Zhu et al., 2006, p. 601). The corresponding SSWs identified (e.g., ‘interactive,’ ‘channel,’ ‘push,’ ‘ubiquitous’) substantiate this view, making the term generic and outdated, since any contemporary (digital) platform is based on the premises of internet technologies. Thus, we argue that the term has become obsolete.

In sum, our data provides evidence that research on platforms as digital tools considers diverse layers of technologies, comprising hardware such as mobile devices, computing infrastructures, protocols and networking technologies, software development frameworks, and software execution environments. While it is possible to address each layer specifically, the established set of terms is overlapping. We interpret these overlaps as part of a broader trend in which particular layers of technology are abstracted in favour of considering complete technology stacks.

Online communities and social platforms

The third cluster includes the terms ‘social platform,’ ‘social media platform,’ ‘online platform,’ ‘social networking platform,’ and ‘communication platform,’ exhibiting a maximum distance of 0.6, showing the considerable degree of semantic overlap.

‘Social platform’ as a concept lacks a popular and well-cited definition in IS research, although it is mainly seen as a tool for connecting users to enable interaction and communication (Cheung et al., 2011; Mitchell-Wong et al., 2008). Examples of ‘social platforms’ include social networking websites, online discussion forums, and blogs (Cheung et al., 2014; Cui et al., 2016), thus covering a wide range of such platforms. We choose ‘social platforms’ to represent the cluster itself as an umbrella term covering different platforms that focus on social interactions and connections.

The concept of ‘social media platform’ has evolved to become the most frequently used platform term in the IS knowledge base (cf. Table 2). ‘Social media platforms’ “facilitate information exchange between users” (Kaplan & Haenlein, 2010, p. 60), covering different forms of user-generated content (Wade et al., 2020). Kallinikos and Constantinou (2015, p. 73) propose viewing ‘social media platforms’ as “huge interaction machines rather than algorithms.” The SSWs underline these definitions (e.g., ‘microblog,’ ‘twitter,’ ‘channel,’ ‘LinkedIn,’ ‘strategically,’ ‘facebook,’ ‘widespread,’ ‘instagram,’ ‘targeted,’ and ‘commercial.’) A ‘social media platform’ refers to user-generated content and media, and their sharing in online communities. Thus, the term is a specialized version of ‘social platform.’

‘Online platform’ is another prevalent platform term in this cluster. At a first glance, ‘online platform’ seems remotely connected to the other terms in this cluster, since it does not convey clear semantics apart from its reference

³ For more information visit <https://www.eclipse.org/eclipseide/>.

to the Internet. Definitions for ‘online platforms’ are scarce, comprising a technological basis, delivering and aggregating services (Batura et al., 2015), and facilitating involvement and interactions for users and crowds (Nevo & Kotlarsky, 2020; OECD, 2019). The context in which platforms are characterized as ‘online platforms’ is, however, diverse and covers a broad outset of IS research areas focusing on social interactions and online communities, as evidenced by the cluster analysis and the SSWs. Thus, we propose the discontinuation of the terms due to its unclear and inconsistent meaning, and its nonspecific reference to digital platforms. Instead, we advise IS researchers to use ‘digital platform’ (cf. Section 4.5) or to refer to a more specific platform term.

A ‘social networking platform’ links “networks of users with the providers of various services and applications” (Bakos & Katsamakas, 2008, p. 172). ‘Social networking platforms’ often integrate third-party developers, who can integrate third-party content on the platform (Felt & Evans, 2008). In the IS literature and the SSWs we can identify ‘social networking platforms’ being LinkedIn, Yammer, Facebook, and Twitter. As such, ‘social networking platforms’ focus on promoting interactions between users, helping them with networking and connecting activities. Thus, we view ‘social networking platforms’ as another specialization of ‘social platforms.’

The last term in this cluster is ‘communication platform.’ Considering the evolution of platform terms in the history of IS (cf. Figure 4), this term was deemed to represent the archetype of a ‘social platform’ up to 2008, whereas they have now turned into the least popular terms in this cluster. This observation aligns with definitions of a ‘communication platform’ reaching back to 1998 (Bertino, 1998), but it should be noted that the interpretation of ‘communication platform’ has evolved from being a client-server architecture, predominantly on the Internet (Bertino, 1998), to having become a system which enables users to send, read, and reply to direct messages from other Internet users (Chang & Wu, 2014). As such, ‘communication platforms’ can improve knowledge management across organizations by decreasing required human communication (Dullaert et al., 2009; Jin & Kotlarsky, 2012). The SSWs underline this interpretation (e.g., ‘sms,’ ‘intranet,’ ‘messaging,’ ‘channel,’ ‘dialogue,’ ‘connect.’) We, therefore, view ‘communication platforms’ as a third specialization of ‘social platforms.’

Economic platforms as digital markets

An economic view on platforms as digital markets covers the terms ‘crowdfunding platform’ and ‘crowdsourcing platform’. Both types of platforms focus on economic effects and crowd involvement. Due to these terms’

definitions and their similarity to (two-/multi-) ‘sided platforms’, we incorporated ‘sided platform’ in this cluster, to comprise three terms.

‘Crowdfunding platforms’ enable the “financing of a project or a venture by a group of individuals instead of professional parties” (Schwienbacher & Larralde, 2010, p. 370) via internet services (Burtch et al., 2013; Liu et al., 2015). As such, ‘crowdfunding platforms’ allow individuals to freely present ideas to an online community to raise financial support for the realization of their products or services, matching ideas with investors (Gerber et al., 2012). Examples of ‘crowdfunding platforms’ are RocketHub, Kickstarter, and IndieGoGo (Gerber et al., 2012), which also feature at the top of the list of the term’s SSWs.

Crowdsourcing is defined as a “type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task” (Estellés-Arolas & González-Ladrón-de-Guevara, 2012, p. 197). ‘Crowdsourcing platforms’ enable these activities for online communities, propelling rivalry by incentives (Bauer et al., 2016), which is supported by the SSWs (e.g., ‘brokerage,’ ‘intermediary,’ ‘advice,’ and ‘challenge’). Thus, while both terms refer to digital markets, ‘crowdsourcing platforms’ and ‘crowdfunding platforms’ differ by the type of activities they enable and the purpose of the community involvement.

A closer inspection of the papers containing ‘sided platform’ reveals that this term encapsulates more specific terms, such as ‘multi-sided platform,’ and ‘two-sided platform,’ which were automatically shortened during data pre-processing, since numbers are function words. Multi-sided platforms “coordinate the demand of distinct groups of customers who need each other in some way” (Evans, 2003, p. 325), e.g., in dating clubs and yellow pages. Two-sided platforms focus on (in-)direct network effects between two sides of a market, mostly investigating pricing structures (Hagiu, 2007). Research on both two-sided and multi-sided platforms focusses on economic effects, e.g., pricing, economic models, and market competition (Evans, 2003; Hagiu, 2007; Hagiu & Wright, 2015) also supported by the SSWs (e.g., ‘champion,’ ‘cryptocurrency,’ ‘instruments,’ and ‘enabler.’)

All terms comprising this cluster focus on investigating the economic effects on platforms. This observation is consistent with the term’s evolution over time (cf. Figure 3). While all terms emerged after 2010, they have become increasingly popular. Interestingly, they refer to different types of platforms, such that they can co-exist without the substantial overlay that we identified in other clusters (e.g., clusters 1 and 2).

General properties of platforms as IS artifacts for value co-creation

We identified the fifth cluster to cover terms that refer to platforms from an abstract perspective, highlighting the general properties that constitute the inner core of platforms and hold across diverse application scenarios.

‘Digital platforms’ refer to platforms as IS artifacts designed to attract and incorporate content supplied by third parties. A prominent review of research on ‘digital platforms’ is provided by de Reuver et al. (2018), differentiating a technical perspective on ‘digital platforms’ as “an extensible codebase to which complementary third-party modules can be added” from a socio-technical view of ‘digital platforms’ as “technical elements (of software and hardware) and associated organizational processes and standards” (de Reuver et al., 2018, p. 127). Both definitions refer closely to earlier conceptualizations of platforms (Sedera et al., 2016; Tiwana et al., 2010) and treat an extensible codebase as an indispensable feature constituting ‘digital platforms.’ In contrast with this technical viewpoint, the SSWs associated with ‘digital platform’ in our dataset show the evolution of the term which now refers to platforms in a much broader sense, comprising strategic (e.g., tactic, ambidextrous, strategizing), economic (e.g., intermediary, payment, marketplace), organizational (e.g., meta-organization, ecosystem, startup), and technological (e.g., blockchain, architecture, IoT) aspects. Against this backdrop, ‘digital platforms’—the platform term displaying the most significant growth rate in research papers covering recent years—seem to have become the basic concept in IS when referring to platforms in an abstract sense, highlighting the platforms’ most central properties, while abstracting from specific aspects of their design and use.

A similarly abstract view on ‘digital platforms’ is reflected in the term ‘service platform.’ A ‘service platform’ is a “modular structure that consists of tangible and intangible components (resources) and facilitates the interaction of actors and resources (or resource bundles)” (Lusch & Nambisan, 2015, p. 162). Thus, the concept focuses on the role of a platform to facilitate value co-creation among the actors interacting on the platform, thereby constituting a service ecosystem (Lusch & Nambisan, 2015). Service—as viewed from the service-dominant logic standpoint (Vargo & Lusch, 2004) on which this definition is based—is an abstract concept, referring to “the application of specialized competences [...] through deeds, processes, and performances for the benefit of another entity or the entity itself” (Vargo & Lusch, 2004, p. 2). While particular types of service can be co-created with ‘digital platforms,’ ‘service platforms’ as a concept deals with the mechanisms that lead to co-creating value irrespective of these more specialized platform types. For instance, a ‘service platform’ can explain the

mechanisms constituting value co-creation on a multi-sided market or a social media platform by applying an abstract lens on value co-creation, even if the way in which value co-creation works on each platform type might differ.

While a ‘cloud platform’ lacks a consistent definition in the extant literature, its recurring properties concern technical and business model aspects in terms of how the platform “deploy[s] software via the Internet” (Katzan, 2009, p. 256). A boundary-spanning role—integrating technical aspects and business aspects—is reflected in the term’s position in our dendrogram, and by its SSWs (business aspects include ‘marketplace,’ ‘consulting,’ ‘commercialize,’ ‘brokerage,’ ‘competency,’ technical aspects include ‘virtualization,’ ‘proprietary,’ ‘infrastructure,’ ‘tenancy,’ and multiple classes of applications such as enterprise resource planning or customer relationship management). Consistent with this boundary-spanning role, Katzan (2009, p. 260) defines a ‘cloud platform’ as “an operating system that runs in the cloud and supports the software-as-a-service concept.” From a technical point of view, a ‘cloud platform’ provides infrastructure (infrastructure-as-a-service, e.g., virtualized hardware), development platforms for software (platform-as-a-service, e.g., middleware), or software (software-as-a-service, e.g., business applications) as a shared pool of virtualized resources that is scalable and available. Third-party users can deploy their software and have it hosted in the cloud as a managed service, or they can deploy and run their own applications in the cloud (Katzan, 2009). A prominent example is Google Cloud Platform,⁴ which supplies business customers with, amongst others, solutions to design cloud-based data storage, monitoring and analytics solutions, mobile apps, and media solutions. As one aspect of cloud computing, ‘cloud platforms’ refer to as-a-service business models that build on metered or subscription pricing models, depending on the resources consumed by users over time (Katzan, 2009). As-a-service business models help users, amongst others, to adjust their computing resources to flexible demand using pooled resources, to lower their capital lockup, and to have flexible and scalable access to computing resources that are placed “somewhere in the Internet”, i.e., in the cloud (Katzan, 2009, p. 257).

Sharing platforms

In our cluster analysis, the term ‘sharing platform’ shows the highest distances (> 0.8) from all other clusters, marking it out as a cluster in its own right. At a higher level of detail our data indicate that the vocabulary associated with ‘sharing platform’ relates to both platform economics (cf.

⁴ For more information visit <https://cloud.google.com/solutions/manufacturing>.

Section 4.4) and to online communities (cf. Section 4.3), displaying a strong connection with economic transactions (SSW: ‘rental,’ ‘barter,’ ‘lending,’ ‘rideshare,’ ‘commercial,’ ‘money,’ ‘carsharing’) and peer-to-peer interactions, as frequently implemented on ‘sharing platforms’ (SSW: ‘P2P,’ ‘C2C,’ ‘ecosystems’).

Despite this apparent overlap of social and economic aspects in sharing, the IS literature uses ‘sharing platform’ differently from economic platforms or online communities. By definition, ‘sharing platforms’ “facilitate sharing among people who do not know each other, and who lack friends or connections in common, [...] mak[ing] stranger sharing less risky and more appealing because they source information on users via the use of ratings and reputations” (Frenken & Schor, 2017, p. 4). As such, they provide a mediating technology to enable sharing between different parties (Sutherland & Jarrahi, 2018). Seminal work on sharing (Belk, 2010) outlined why sharing, gift giving, and marketplace exchange represent very different prototypes of interactions, emphasizing that “sharing tends to be a communal act that links us to other people” (Belk, 2010, p. 717), and which can be seen as “nonreciprocal pro-social behavior” (Benkler, 2004, p. 275). Caring for others and social bonding are seen as defining characteristics of sharing (Belk, 2010). The same characteristics can also be found in research on online communities; Karahanna et al. (2018) identify relatedness as a psychological need in the social media context, defined as the “need to interact, be connected to, and experience caring for others” (Karahanna et al., 2018, p. 740). In contrast to sharing, economic exchange is traditionally characterized by a transfer of ownership and an impersonal relationship between exchanging parties (Belk, 2010). While the Internet has brought up new forms of Internet-facilitated sharing (Belk, 2014)—e.g., enabling people to share data with strangers and material goods in their neighborhoods—sharing is still supposed to be a social, not-for-profit interaction.

However, on digital platforms, references to ‘sharing’ are incongruent with its original meaning, rather pointing to a collaborative consumption of under-utilized resources. Collaborative consumption occurs when “people coordinat[e] the acquisition and distribution of a resource for a fee or other compensation,” also termed pseudo-sharing “in that they often take on a vocabulary of sharing (e.g., ‘car sharing’), but are more accurately short-term rental activities” (both: Belk, 2014, p. 1597). For instance, the platforms featured in the SSWs of ‘sharing platform’ (Airbnb, BlaBlaCar, or Lyft) enable users to consume houses or transportation collaboratively, using digital platforms that enable impersonal interactions among service providers and customers on a digital multi-sided market.

We conclude that as a theoretical lens, proper sharing leans towards the use of platforms to establish social communities, while collaborative consumption leans towards

market-based interactions as a domain of platform economics. We conclude that the concept of a ‘sharing platform’ is often used inconsistently with foundational concepts of sharing and, therefore, cannot provide the missing link to connect “social network research, such as research on collective intelligence, with the domain of online social commerce as it is established in C2C interactions” (Puschmann & Alt, 2016, p. 95). We propose that—instead of using ‘sharing platform’—digital platforms that link actors for the purpose of nonreciprocal social behavior should be referred to as ‘social platforms’, whereas the term ‘(two-/multi-) sided platforms’ should be used to refer to a digital platform that enables impersonal market-based interactions, including collaborative consumption in a peer-to-peer network. With this distinction, research can avoid confusion concerning the terms ‘sharing platform’ and sharing.

Decomposed lexicon of platforms in IS

Based on the interpretation and discussion of the terms we discovered in our data-driven study, we strive to propose a consolidated lexicon of platform terms in IS research that reduces complexity and enables future research to study the phenomenon of digital platforms and their subtypes in detail. We built the lexicon by hierarchically decomposing the knowledge base on digital platforms in IS. Decomposition is a well-known and first-hand solution to break down complex structures (Alexander, 1964), like the rhizome structure currently appertaining to the IS knowledge base on digital platforms. While there is always more than one solution to decompose a complex structure or problem, every single one is always more suitable for a particular purpose (Alexander, 1964). The main goal of decomposition is to identify subsets “whose *internal* interactions are very rich” (Alexander, 1964, p. 124) with “as little interaction *between* subsets as possible” (ibid). At the same time, it is important to acknowledge that underemphasizing the relationships between the subsets of a decomposed structure is a critical issue that needs to be prevented (Alexander, 1964). Building on decomposition, we aim to develop a lexicon of digital platforms in IS, which enables future research to address specific terms and concepts better than today’s rhizomatic structure of platform terms.

A decomposed model of a structure enables the study of not only the details of the system but also the interactions between all parts of the system as a whole (Simon, 1996). Thus, our decomposed lexicon of digital platforms in the IS literature provides researchers with a systematization that can be applied to study digital platforms and related phenomena from different (isolated) theoretical perspectives and by combining multiple perspectives to study a phenomenon at a more abstract level or as a whole (Fig. 5). They might

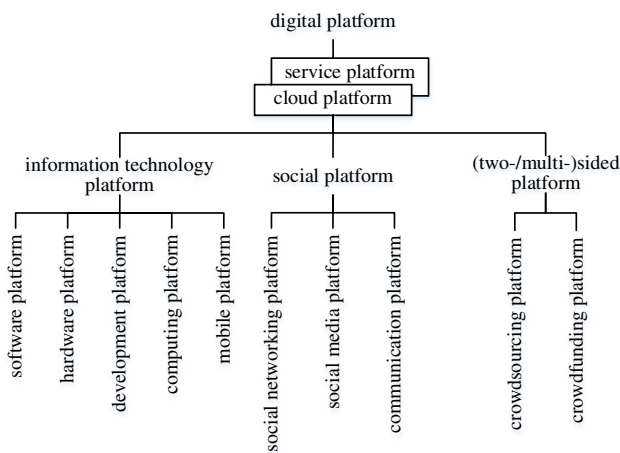


Fig. 5 A decomposed model of terms that can be used or combined as lenses to study platforms in IS research

either decide to use one particular concept to focus on, or they might combine different concepts to investigate more complex phenomena. In this regard, the root node of our model, ‘digital platform’ comprises all features that relate to digital platforms and are decomposed in the subsets. IS artifacts exhibit technical and social properties. A ‘digital platform’ as a whole is, thus, a generative IS artifact that provides a mutual core of technology and organizational arrangements, inviting compatible and complementary resources (e.g., hardware, software, or content) from third parties to enable the emergence of digital online communities or markets (de Reuver et al., 2018).

Building on digital platforms, we posit that two overarching views can be applied when studying platforms. First, ‘service platforms’ point to the role that platforms play for co-creating value among the stakeholders in service (eco-) systems, such as service providers (including a platform owner and platform provider) and service customers. In this sense, a ‘service platform’ is a view on digital platforms as structures that establish value co-creation in service ecosystems, enabling actors to provide, access, and integrate complementary resources in service-for-service exchanges, building on a mutual core of technology and organizational arrangements (Lusch & Nambisan, 2015; Vargo & Lusch, 2004). Second, from a more technical perspective, ‘cloud platforms’ refer to digital platforms as IT artifacts that ought to be designed in specific ways as a prerequisite to enable actors to co-create value. ‘Cloud platforms’ provide an abstract view on an operating system that runs in the cloud and provides a shared pool of virtualized, scalable and available resources in the form of infrastructure, development platforms for software, or software. Both views are valid since cloud platforms are strongly related to as-a-service business models, as evidenced by their SSWs. For this reason, they provide complementary, yet abstract,

views on the design, form, and function of platforms as IT artifacts (cloud platforms), and on the co-creation of value established through users interacting with a digital platform (service platform). Due to their role as views, neither of these platforms concepts provides a direct super-structure for the more detailed platform concepts discussed subsequently, which aligns with the purpose of a taxonomy, provided the ‘subclass of’ primitive is not applied to the edge connecting these concepts (Gomez-Perez & Corcho, 2002).

‘Information technology platforms,’ ‘social platforms,’ and ‘(two-/multi-)sided platforms’ are three more detailed perspectives on digital platforms. Triangulating the insights identified from clusters one and two, the term ‘information technology platform’ is identified as an overarching concept referring to technical views on platforms as IT artifacts. All other terms in clusters one and two are outdated (cluster one: ‘technical platform,’ and ‘ecommerce platform;’ cluster two: ‘internet platform’), nonspecific (cluster one: ‘technology platform,’ and ‘technological platform’), refer to a platform characteristic instead of a platform type (cluster two: ‘open platform’ refers to platform openness), or represent subclasses of ‘information technology platforms’ (cluster two: ‘software platform,’ ‘hardware platform,’ ‘development platform,’ ‘computing platform,’ and ‘mobile platform’). To provide a consistent decomposition of platform terms, we conclude that the state-of-the-art of platform research refers to the platform as an IT artifact as an ‘information technology (IT) platform’ (Table 4).

Combining the insights from the five platform terms in the third cluster, we identified that a ‘social platform’ is a generalization of ‘social media platform,’ ‘social networking platform,’ and ‘communication platform.’ These three terms differ by focusing on direct message exchange and communication (communication platform), creating and sharing user-generated content (social media platform), or networking and direct interactions (social networking platform).

Still, all three terms refer to social interactions that might often take place on the same social platform. Facebook, for instance, implements multiple features of ‘communication platforms’ (e.g., sending messages to peers), ‘social networking platforms’ (e.g., connect with friends and join groups), and ‘social media platforms’ (e.g., enabling influencers to share user-generated content, advertise products, and sell services). We advise researchers to use a more specific platform term when outlining their research, or use the broader term ‘social platform’ when referring to digital platforms that enable social interactions.

Concerning platform terms viewed from an economic viewpoint, our analysis identified ‘crowdfunding platform,’ ‘crowdsourcing platform,’ and ‘sided platform.’ As both crowdfunding and crowdsourcing platforms can be two-sided or multi-sided platforms, we advocate ‘(two-/multi-) sided platform’ to be a more general term, specializing into

Table 4 A lexicon of significant platform terms for IS research

Terms	Definitions	Selected References
digital platform	Generative IS artifacts that provide a mutual core of technology and organizational arrangements, inviting compatible and complementary resources (e.g., hardware, software, or content) from third-parties, to enable the emergence of digital online communities or markets.	de Reuver et al., 2018
service platform	View on digital platforms as structures that establish value co-creation in service ecosystems, enabling actors to provide, access, and integrate complementary resources in service-for-service exchanges, building on a mutual core of technology and organizational arrangements.	Lusch & Nambisan, 2015; Vargo & Lusch, 2004, 2016
cloud platform	Abstracted view on an operating system that runs in the cloud and provides infrastructure, development platforms for software, or software as a shared pool of virtualized resources that is scalable and available.	Katzan, 2009
information technology platform	Digital platform as a conglomerate of adaptable information technology components that interoperate via interfaces and compose a layered, modular architecture.	Kraemer & Dedrick, 2002; Tiwana et al., 2010; Yoo et al., 2010
software platform	Information technology platform that provides software packages, and application systems.	Taudes et al., 2000
hardware platform	Information technology platform that composes a set of information technology hardware components that execute software applications and enables their re-use.	de Michell & Gupta, 1997; Keutzer et al., 2000
development platform	Information technology platform that provides integrated development environments comprising frameworks and tools for software development and execution.	Mehra et al., 2011
computing platform	Information technology platform that provides computational resources and can be employed to solve computational problems.	Athanas & Abbott, 1995
mobile platform	Information technology platform that employs mobile devices as boundary objects to set up platform ecosystems subject to control mechanisms and generativity.	Ghazawneh & Henfridsson, 2013; Sørensen et al., 2015; Tilson et al., 2010
social platform	Digital platform for connecting different groups of users to enable interactions and communication between users.	C. M. Cheung et al., 2011; C. M. Cheung et al., 2014; Cui et al., 2016; Mitchell-Wong et al., 2008
social media platform	Social platform facilitating exchange and sharing of user-generated content in online communities.	Kallinikos & Constantiou, 2015; Kaplan & Haenlein, 2010; Wade et al., 2020
social networking platform	Social platform enabling direct interactions between users and supporting networking and connecting activities.	Bakos & Katsamakas, 2008; Felt & Evans, 2008
communication platform	Social platform for direct message exchange between users to improve knowledge management and integration.	Bertino, 1998; Dullaert et al., 2009; Jin & Kotlarsky, 2012
(two-/multi-) sided platform	Digital platform focusing on (in-)direct network effects between (two / multiple) sides of a market, coordinating the demand of its distinct groups of customers.	Evans, 2003; Hagiu, 2007; Hagiu & Wright, 2015
crowdfunding platform	(Two-/multi-) sided platform enabling the financing of an idea, a project, or a venture from (groups of) individuals by investors.	Burtch et al., 2013; Gerber et al., 2012; Schwienbacher & Larralde, 2010

Table 4 (continued)

Terms	Definitions	Selected References
crowdsourcing platform	(Two-/multi-) sided platform enabling online communities to match tasks to individuals of varying knowledge, heterogeneity, and number in open calls, propelling rivalry by the (mostly monetary) incentives.	Bauer et al., 2016; Estellés-Arolas & González-Ladrón-de-Guevara, 2012

either ‘crowdfunding platform’—for multi-sided platforms to raise funds—or ‘crowdsourcing platform’—for multi-sided platforms to source work from anonymous online communities.

Conclusion

With our collection, analysis, and interpretation of digital platform terms in IS research, we offer two theoretical contributions. Having identified and analyzed platform terms from a compilation of 11,049 research papers published in leading peer-reviewed IS journals and conference proceedings in the past 44 years, and having applied an inductive text-mining approach using machine learning, our analysis is the first to cover the entire history of platform research in the IS knowledge base. With the identified terms and their SSWs, we contribute much-needed empirical evidence on platform research. Our decomposed model of digital platform terms and our lexicon of platform terms in IS research provide a common baseline to guide future platform research in studying isolated phenomena from a specific theoretical perspective, and to enable investigating broader phenomena on a more abstract level by combining different theoretical perspectives on platforms. At the same time, our analysis might enable researchers to identify and interpret older platform papers that used different terms in the formation phase of platform research. We consider both temporal directions—shaping the future and accessing the past of platform research—as equally important applications of our decomposed model and lexicon. We encourage others to apply the decomposed model of platform terms by, for example, using the term cloud platform when conducting future research on technical facets of digital platforms, while we advocate for using the term service platform when studying actors and services. Additionally, the ability to switch between different views on platforms also enables researchers to study the relationships between the identified subsets of platforms terms. This rationale can be applied to all other elements of the model. Consistently using the terms will help build and refine a consistent lexicon of digital platforms in the IS discipline in the future.

Although we took great precautions with generalizing the results of our data-driven analysis, our study is subject

to common limitations inherent to text mining, the application of machine learning algorithms, and interpretive research. While at first glance a total number of 11,049 papers seems sufficient to justify the application of these methods, the size of our data set and choice of sources might still limit the generalizability of our contribution. We restricted our analysis to top-ranked journals and conference proceedings in the IS discipline (Tarafdar & Davison, 2018). As platforms are a topic of high interest to both researchers and practitioners, an even more inclusive consideration of practice-oriented journals, conference proceedings, or white papers could have led to the identification of additional or different terms. Since the results of the application of unsupervised machine learning are highly dependent on large amounts of data, extending the data set could lead to even better results to inform the manual coding and interpretation processes. On the other hand, extending the data set to include papers of lower scientific quality could also lower the data quality and, therefore, lead to less reliable results. Since our approach was to systematize and constitute the lexicon of platform terms in IS, we decided in favor of scientific excellence, at the expense of considering a slightly reduced but still sizeable data set. Faced with this trade-off, other researchers might take a different decision.

Likewise, data quality leaves its mark on the results of data-driven papers, since some papers might only have been included in the first research step because their bibliographies contained papers that referred to the word ‘platform’ in their titles. Likewise, the second step of our research process could potentially produce sub-optimal results, as authors might not use identical platform terms throughout their papers. To deal with these limitations appropriately, we performed different configurations and verified the quality of the results in each step of our research process for optimal parameter selection. We documented all decisions made in the analysis to enable others to replicate our research process. One important decision made in our analysis was to exclude platform terms that appeared in fewer than 150 papers (Table 2). Since we subsequently developed the decomposed model based on the reduced database, it is conceivable that including more terms would have produced a broader—yet less specific and common—collection of platform terms. A closer inspection of the excluded platform

terms, however, suggests that these terms could still fit into our lexicon.

A general limitation of text mining and data aggregation techniques is that they do not always allow to retrospectively draw conclusions from the data. In our case, this means that we cannot refer to single papers or their context to explain the overall results identified from the dendrogram. To reduce the impact of this limitation on our contribution, we discussed the results of our data analysis with reference to definitions in the existing literature. Further, by conducting the majority of the research process automatically, we aimed to limit the (potentially detrimental) influence of the authors (bias and subjectivity), thus increasing the generalizability of our results, but acknowledge that the lack of human input could limit the conceptual insights gained from the results.

As regards the interpretation of our results, we identified a lack of common definitions for certain platform terms, which somewhat hampers the objective discussion of our results. Additionally, while we are confident that we provide a decomposition of platform terms with maximum inner meaningfulness and minimum interactions between the subsets, others might arrive at different suitable decompositions. However, we do not see this natural bias of decomposition as a weakness of our paper but want to invite others to critically review, revise, and enhance the decomposed model and corresponding lexicon of platforms terms.

There are ample opportunities to validate, discuss, and extend our results through quantitative, qualitative, and design-oriented studies. First, the research process presented here could be replicated in neighboring research disciplines of the IS field (e.g., economics, computer science, service science, marketing, engineering) to compare platform research across these disciplines. This effort would enhance our knowledge about the design, implementation, and evolution of platform concepts. An inter-disciplinary analysis might indicate how platforms can serve as boundary objects that bridge platform-related research carried out in different disciplines. Second, since our research process focused on analyzing textual data with data mining and machine learning, conceptual studies could challenge and extend our lexicon of platforms in IS.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12525-022-00530-6>.

Acknowledgements This paper is part of two research projects. The project DigiBus was funded under the promotion sign 005-1807-0107 by the Ministry of Economics, Innovation, Digitalization, and Energy of the State of North Rhine-Westphalia (MWIDE.NRW). The project FLEMING was funded under the promotion sign 03E16012F by the Federal Ministry for Economic Affairs and Climate Action (BMWK). Furthermore, we acknowledge the work by Philipp Hansmeier, who supported us in conducting a comprehensive literature review documented in the online appendix.

Funding Open Access funding enabled and organized 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

- Alexander, C. (1964). *Notes on the synthesis of form* (Vol. 5). Harvard University Press.
- Almeida, J., Barbosa, L., Pais, A., & Formosinho, S. J. (2007). Improving hierarchical cluster analysis: A new method with outlier detection and automatic clustering. *Chemometrics and Intelligent Laboratory Systems*, 87(2), 208–217. <https://doi.org/10.1016/j.chemolab.2007.01.005>
- Antons, D., & Breidbach, C. F. (2018). Big data, big insights? Advancing service innovation and design with machine learning. *Journal of Service Research*, 21(1), 17–39. <https://doi.org/10.1177/1094670517738373>
- Asadullah, A., Faik, I., & Kankanhalli, A. (2018). Evolution mechanisms for digital platforms: A review and analysis across platform types. *Proceedings of the 39th International Conference on Information Systems (ICIS)*.
- Athanas, P. M., & Abbott, A. L. (1995). Real-time image processing on a custom computing platform. *Computer*, 28(2), 16–25. <https://doi.org/10.1109/2.347995>
- Atkinson, C., & Brooks, L. (2005). In the age of the Humanchine. *Proceedings of the 26th International Conference on Information Systems (ICIS)*.
- Bakos, Y., & Katsamakas, E. (2008). Design and ownership of two-sided networks: Implications for internet platforms. *Journal of Management Information Systems*, 25(2), 171–202. <https://doi.org/10.2753/MIS0742-1222250208>
- Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In A. Gawer (Ed.), *Platforms, markets and innovation*. Edward Elgar. <https://doi.org/10.2139/ssrn.1265155>
- Basole, R. C., & Karla, J. (2011). On the evolution of Mobile platform ecosystem structure and strategy. *Business & Information Systems Engineering*, 3(5), 313–322. <https://doi.org/10.1007/s12599-011-0174-4>
- Batura, O., van Gorp, N., & Larouche, P. (2015). *Online platforms and the EU digital single market: A response to the call for evidence by the house of lord's internal market sub-committee*. E-Conomics. https://ec.europa.eu/information_society/newsroom/image/document/2016-7/nikolai_van_gorp_-_response_e-economics_to_the_uk_house_of_lords_call_for_evidence_14020.pdf. Accessed 22 Feb 2022
- Bauer, J., Franke, N., & Tuertscher, P. (2016). Intellectual property norms in online communities: How user-organized intellectual property regulation supports innovation. *Information Systems Research*, 27(4), 724–750. <https://doi.org/10.1287/isre.2016.0649>

- Belk, R. (2010). Sharing. *Journal of Consumer Research*, 36(5), 715–734. <https://doi.org/10.1086/612649>
- Belk, R. (2014). You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, 67, 1595–1600. <https://doi.org/10.1016/j.jbusres.2013.10.001>
- Benkler, Y. (2004). Sharing nicely: On shareable goods and the emergence of sharing as a modality of economic production. *The Yale Law Journal*, 114(2), 273–358. <https://doi.org/10.2307/4135731>
- Benlian, A., Hilkert, D., & Hess, T. (2015). How open is this platform? The meaning and measurement of platform openness from the Complementers' perspective. *Journal of Information Technology*, 30(3), 209–228. <https://doi.org/10.1057/jit.2015.6>
- Berente, N., Seidel, S., & Safadi, H. (2019). Research commentary—data-driven computationally intensive theory development. *Information Systems Research*, 30(1), 50–64. <https://doi.org/10.1287/isre.2018.0774>
- Bertino, E. (1998). Data security. *Data & Knowledge Engineering*, 25(1–2), 199–216. <https://doi.org/10.1109/CMPSAC.1998.716660>
- Blechar, J., Constantiou, I. D., & Damsgaard, J. (2006). Exploring the influence of reference situations and reference pricing on mobile service user behaviour. *European Journal of Information Systems*, 15(3), 285–291. <https://doi.org/10.1057/palgrave.ejis.3000618>
- Boudreau, K. J. (2010). Open platform strategies and innovation: Granting access vs. devolving control. *Management Science*, 56(10), 1849–1872. <https://doi.org/10.1287/mnsc.1100.1215>
- Boudreau, K. J., & Hagiu, A. (2009). Platform rules: Multi-sided platforms as regulators. In A. Gawer (Ed.), *Platforms, markets and innovation* (pp. 163–191). Edward Elgar. <https://doi.org/10.2139/ssrn.1269966>
- Burtch, G., Ghose, A., & Wattal, S. (2013). An empirical examination of the antecedents and consequences of contribution patterns in crowd-funded markets. *Information Systems Research*, 24(3), 499–519. <https://doi.org/10.1287/isre.1120.0468>
- Butler, B. S., Bateman, P. J., Gray, P. H., & Diamant, E. I. (2014). An attraction-selection-attrition theory of online community size and resilience. *MIS Quarterly*, 38(3), 699–728. <https://doi.org/10.25300/MISQ/2014/38.3.04>
- Chang, H. H., & Wu, L. H. (2014). An examination of negative e-WOM adoption: Brand commitment as a moderator. *Decision Support Systems*, 59(2), 206–218. <https://doi.org/10.1016/j.dss.2013.11.008>
- Cheung, C. M., Lee, M. K. O., & Jin, X.-L. (2011). Customer engagement in an online social platform: A conceptual model and scale development. *Proceedings of the 32nd International Conference on Information Systems (ICIS)*.
- Cheung, C. M., Xiao, B. S., & Liu, I. L. (2014). Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. *Decision Support Systems*, 65(2), 50–58. <https://doi.org/10.1016/j.dss.2014.05.002>
- Churilov, L., Bagirov, A., Schwartz, D., Smith, K., & Dally, M. (2005). Data mining with combined use of optimization techniques and self-organizing maps for improving risk grouping rules: Application to prostate Cancer patients. *Journal of Management Information Systems*, 21(4), 85–100. <https://doi.org/10.1080/07421222.2005.11045826>
- Constantinides, P., Henfridsson, O., & Parker, G. G. (2018). Introduction—platforms and infrastructures in the digital age. *Information Systems Research*, 29(2), 381–400. <https://doi.org/10.1287/isre.2018.0794>
- Cui, X., Lai, V. S., & Lowry, P. B. (2016). How do bidders' organism reactions mediate auction stimuli and bidder loyalty in online auctions? The case of Taobao in China. *Information & Management*, 53(5), 609–624. <https://doi.org/10.1016/j.im.2016.01.005>
- de Michell, G., & Gupta, R. K. (1997). Hardware/software co-design. *Proceedings of the IEEE*, 85(3), 349–365. <https://doi.org/10.1109/5.558708>
- de Reuver, M., Sørensen, C., & Basole, R. (2018). The digital platform: A research agenda. *Journal of Information Technology*, 33(2), 124–135. <https://doi.org/10.1057/s41265-016-0033-3>
- Debortoli, S., Müller, O., Junglas, I., & vom Brocke, J. (2016). Text Mining for Information Systems Researchers: An annotated topic modeling tutorial. *Communications of the Association for Information Systems*, 39, 110–135. <https://doi.org/10.17705/ICAIS.03907>
- Deleuze, G., & Guattari, F. (1979). *Rhizome: Introduction*. Les Éditions de Minuit.
- Dullaert, W., Neutens, T., Vanden Berghe, G., Vermeulen, T., Vernimmen, B., & Witlox, F. (2009). MamMoeT: An intelligent agent-based communication support platform for multimodal transport. *Expert Systems with Applications*, 36(7), 10280–10287. <https://doi.org/10.1016/j.eswa.2009.01.049>
- Eaton, B. D., Elaluf-Calderwood, S., Sørensen, C. and Yoo, Y. (2015). Distributed tuning of boundary resources: the case of apple's iOS service system. *MIS Quarterly: Special Issue on Service Innovation in a Digital Age*, 39(1): 217–243. <https://doi.org/10.25300/MISQ/2015/39.1.10>
- Elbanna, A., & Linderot, H. C. J. (2015). The formation of technology mental models: The case of voluntary use of technology in organizational setting. *Information Systems Frontiers*, 17(1), 95–108. <https://doi.org/10.1007/s10796-014-9513-6>
- Estellés-Arolas, E., & González-Ladrón-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200. <https://doi.org/10.1177/0165551512437638>
- Evans, D. S. (2003). The antitrust economics of multi-sided platform markets. *Yale Journal on Regulation*, 20, 325–379. <https://doi.org/10.2139/ssrn.332022>
- Fan, W., Wallace, L., Rich, S., & Zhang, Z. (2006). Tapping the power of text mining. *Communications of the ACM*, 49(9), 76–82. <https://doi.org/10.1145/1151030.1151032>
- Fan, L.-S., Zeng, L., & Luo, S. (2015). Chemical-looping technology platform. *AIChE Journal*, 61(1), 2–22. <https://doi.org/10.1002/aic.14695>
- Felt, A., & Evans, D. (2008). Privacy protection for social networking platforms. *Workshop on Web 2.0 Security and Privacy, USA*, 1–8.
- Frawley, W. J., Piatetsky-Shapiro, G., & Matheus, C. J. (1992). Knowledge discovery in databases: An overview. *AI Magazine*, 13(3), 57–70. <https://doi.org/10.1609/aimag.v13i3.1011>
- Frenken, K., & Schor, J. (2017). Putting the sharing economy into perspective. *Environmental Innovation and Societal Transitions*, 23, 3–10. <https://doi.org/10.4337/9781788117814.00017>
- Fu, W., Wang, Q., & Zhao, X. (2018). Platform-based service innovation and system design: A literature review. *Industrial Management & Data Systems*, 118(5), 946–974. <https://doi.org/10.1108/IMDS-03-2017-0129>
- Gachet, A., & Brézillon, P. (2005). A multi-level model. *Journal of Decision Systems*, 14(1–2), 9–37. <https://doi.org/10.3166/jds.14.9-37>
- Gal-Or, E., Gal-Or, R., & Penmetsa, N. (2018). The role of user privacy concerns in shaping competition among platforms. *Information Systems Research*, 29(3), 698–722. <https://doi.org/10.1287/isre.2017.0730>
- Gawer, A. (Ed.). (2009). *Platforms, markets and innovation*. Edward Elgar. <https://doi.org/10.4337/9781849803311>
- Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. *Research Policy*, 43(7), 1239–1249. <https://doi.org/10.5465/ambpp.2014.278>

- Gerber, E. M., Hui, J. S., & Kuo, P.-Y. (2012). Crowdfunding: Why people are motivated to post and fund projects on crowdfunding platforms. *Proceedings of the International Workshop on Design, Influence, and Social Technologies: Techniques, Impacts, and Ethics*, 2(11), 1–9.
- 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. <https://doi.org/10.1111/j.1365-2575.2012.00406.x>
- Gillespie, T. (2010). The politics of ‘platforms’. *New Media & Society*, 12(3), 347–364. <https://doi.org/10.1002/9781118321607.ch28>
- Goldberg, Y., & Levy, O. (2014). *word2vec Explained: deriving Mikolov et al.'s negative-sampling word-embedding method*. <http://arxiv.org/pdf/1402.3722v1>. Accessed 22 Feb 2022
- Gomez-Perez, A., & Corcho, O. (2002). Ontology languages for the semantic web. *IEEE Intelligent Systems*, 17(1), 54–60. <https://doi.org/10.1109/5254.988453>
- Google Inc. (2013). *Google code archive—word2vec: Tool for computing continuous distributed representations of words*. <https://code.google.com/archive/p/word2vec/>. Accessed 17 Apr 2019
- Gupta, V., & Lehal, G. S. (2009). A survey of text mining techniques and applications. *Journal of emerging technologies in web. Intelligence*, 1(1). <https://doi.org/10.4304/jetwi.1.1.60-76>
- Habermas, J. (2014). *Truth and justification* (1st ed.). Polity Press.
- Hagiu, A. (2007). Merchant or two-sided platform? *Review of Network Economics*, 6(2). <https://doi.org/10.2202/1446-9022.1113>
- Hagiu, A., & Wright, J. (2015). Multi-sided platforms. *International Journal of Industrial Organization*, 43(1), 162–174. <https://doi.org/10.1016/j.ijindorg.2015.03.003>
- Hein, A., Böhm, M., & Krcmar, H. (2018). Platform configurations within information systems research: A literature review on the example of IoT platforms. *Proceedings of the Multikonferenz Wirtschaftsinformatik 2018*.
- Hennig-Thurau, T., Walsh, G., & Schrader, U. (2004). VHB-JOURQUAL: Ein Ranking von betriebswirtschaftlich-relevanten Zeitschriften auf der Grundlage von Expertenurteilen. *Schmalenbachs Zeitschrift Für Betriebswirtschaftliche Forschung*, 56(6), 520–545. <https://doi.org/10.1007/BF03372748>
- Huang, M., Zhao, Y., & Zhu, L. (2011). Research for e-commerce platform security framework based on SOA. *4th International Conference on Biomedical Engineering and Informatics (BMEI)* (pp. 2171–2174). IEEE. <https://doi.org/10.1109/BMEI.2011.6098737>
- Humphreys, P. (2021). Socialising the decision-making process: Transaction provenance decision support. *Journal of Decision Systems*, 1–15. <https://doi.org/10.1080/12460125.2020.1868653>
- Iansiti, M., & Lakhani, K. R. (2014). Digital ubiquity: How connections, sensors, and data are revolutionizing business. *Harvard Business Review*, 92(11), 90–99. <https://doi.org/10.2469/dig.v45.n2.8>
- Iivari, N., Kinnula, M., Molin-Juustila, T., & Kuure, L. (2017). Multiple voices in the maker movement—A Nexus analytic literature review on children, education and making. *Proceedings of the 25th European Conference on Information Systems (ECIS)*.
- Indulska, M., Hovorka, D. S., & Recker, J. (2012). Quantitative approaches to content analysis: Identifying conceptual drift across publication outlets. *European Journal of Information Systems*, 21(1), 49–69. <https://doi.org/10.1057/ejis.2011.37>
- Jin, X., & Kotlarsky, J. (2012). A conceptual framework of knowledge integration in multisourcing arrangements. *Proceedings of the 33rd International Conference on Information Systems (ICIS)*.
- Kallinikos, J., & Constantiou, I. D. (2015). Big data revisited: A rejoinder. *Journal of Information Technology*, 30(1), 70–74. <https://doi.org/10.1057/jit.2014.36>
- Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons*, 53(1), 59–68. <https://doi.org/10.1016/j.bushor.2009.09.003>
- Karahanna, E., Xu, S. X., Xu, Y., & Zhang, A. (2018). The needs-affordances-features perspective for the use of social media. *MIS Quarterly*, 42(3), 737–756. <https://doi.org/10.25300/MISQ/2018/11492>
- Katzan, H. (2009). Cloud software service: Concepts, technology, economics. *Service Science*, 1(4), 256–269. <https://doi.org/10.1287/serv.1.4.256>
- Kazan, E., & Damsgaard, J. (2016). Towards A New Market entry framework for digital payment platforms. *Communications of the Association for Information Systems*, 38, 761–783. <https://doi.org/10.17705/1CAIS.03837>
- Keutzer, K., Newton, A. R., Rabaey, J. M., & Sangiovanni-Vincentelli, A. (2000). System-level design: Orthogonalization of concerns and platform-based design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 19(12), 1523–1543. <https://doi.org/10.1109/43.898830>
- Kraemer, K. L., & Dedrick, J. (2002). Strategic use of the internet and e-commerce: Cisco Systems. *The Journal of Strategic Information Systems*, 11(1), 5–29. [https://doi.org/10.1016/S0963-8687\(01\)00056-7](https://doi.org/10.1016/S0963-8687(01)00056-7)
- Langlois, R. N. (2002). Modularity in technology and organization. *Journal of Economic Behavior & Organization*, 49(1), 19–37. [https://doi.org/10.1016/S0167-2681\(02\)00056-2](https://doi.org/10.1016/S0167-2681(02)00056-2)
- Leclercq-Vandelannoitte, A., Isaac, H., & Kalika, M. (2014). Mobile information systems and organisational control: Beyond the panopticon metaphor? *European Journal of Information Systems*, 23(5), 543–557. <https://doi.org/10.1057/ejis.2014.11>
- Lilleberg, J., Zhu, Y., & Zhang, Y. (2015). Support vector machines and Word2vec for text classification with semantic features. *Proceedings of the IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing*.
- Liu, J., Yang, L., Wang, Z., & Hahn, J. (2015). Winner takes all? The "blockbuster effect" in crowdfunding platforms. *Proceedings of the 36th International Conference on Information Systems (ICIS)*.
- Lusch, R. F., & Nambisan, S. (2015). Service innovation: A service-dominant logic perspective. *MIS Quarterly*, 39(1), 155–176. <https://doi.org/10.25300/MISQ/2015/39.1.07>
- Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
- Márton, A. (2021). Steps toward a digital ecology: Ecological principles for the study of digital ecosystems. *Journal of Information Technology*, 1–16. <https://doi.org/10.1177/02683962211043222>
- McIntyre, D., & Srinivasan, A. (2017). Networks, platforms, and strategy: Emerging views and next steps. *Strategic Management Journal*, 38(1), 141–160. <https://doi.org/10.1002/smj.2596>
- Mehra, A., Dewan, R., & Freimer, M. (2011). Firms as incubators of open-source software. *Information Systems Research*, 22(1), 22–38. <https://doi.org/10.2139/ssrn.1121582>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. <http://arxiv.org/pdf/1301.3781v3>. Accessed 22 Feb 2022
- Miner, G., Delen, D., Elder, J. F., Fast, A., Hill, T., Nisbet, R., Thompson, J., Foley, R., Waner, A., Winters- Miner, L., & Balakrishnan, K. (2012). *Practical text mining and statistical analysis for non-structured text data applications*. Elsevier Academic Press. <https://doi.org/10.1016/C2010-0-66188-8>
- Mitchell-Wong, J., Goh, S. K., Chhetri, M. B., Kowalczyk, R., & Vo, B. Q. (2008). Framework for open, distributed and self-managed social platforms. In L. M. Camarinha-Matos & W. Picard (Eds.), *IFIP – The International Federation for Information Processing. Pervasive collaborative networks* (Vol. 283, pp. 361–368). Springer US. https://doi.org/10.1007/978-0-387-84837-2_37

- Müller, O., Junglas, I., vom Brocke, J., & Debortoli, S. (2016). Utilizing big data analytics for information systems research: Challenges, promises and guidelines. *European Journal of Information Systems*, 25(4), 289–302. <https://doi.org/10.1057/ejis.2016.2>
- Munková, D., Munk, M., & Vozár, M. (2013). Data pre-processing evaluation for text mining: Transaction/sequence model. *Proceedia Computer Science*, 18, 1198–1207. <https://doi.org/10.1016/j.procs.2013.05.286>
- Mustonen-Ollila, E., & Lyytinen, K. (2003). Why organizations adopt information system process innovations: A longitudinal study using diffusion of innovation theory. *Information Systems Journal*, 13(3), 275–297. <https://doi.org/10.1046/j.1365-2575.2003.00141.x>
- Nevo, D., & Kotlarsky, J. (2020). Crowdsourcing as a strategic IS sourcing phenomenon: Critical review and insights for future research. *The Journal of Strategic Information Systems*, 29(4), 101593. <https://doi.org/10.2139/ssrn.3663509>
- Njenga, K., & Brown, I. (2012). Conceptualising improvisation in information systems security. *European Journal of Information Systems*, 21(6), 592–607. <https://doi.org/10.1057/ejis.2012.3>
- Nolas, S.-M. (2008). Supporting a post-merger organization through the practice of research. *Journal of Decision Systems*, 17(1), 79–96. <https://doi.org/10.3166/jds.17.97-117>
- OECD. (2019). *Introduction to online platforms and their role in the digital transformation*. OECD Publishing. <https://doi.org/10.1787/53e5f593-en>
- Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature I. *Libri*, 46(3), 102. <https://doi.org/10.1515/libr.1996.46.3.149>
- Ozer, G., & Anderson, E. (2015). Innovation and breaching strategies in multi-sided platform markets: Insights from a simulation study. *Proceedings of the International Conference on Information Systems (ICIS)*.
- Park, J., Shin, S. K., & Sanders, G. L. (2007). Impact of international information technology transfer on National Productivity. *Information Systems Research*, 18(1), 86–102. <https://doi.org/10.1287/isre.1070.0116>
- Parker, G., van Alstyne, M., & Jiang, X. (2017). Platform ecosystems: How developers invert the firm. *MIS Quarterly*, 41(1), 255–266. <https://doi.org/10.2139/ssrn.2861574>
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54(1), 547–577. <https://doi.org/10.1146/annurev.psych.54.101601.145041>
- Prechelt, L. (2011). Plat_Forms: A web development platform comparison by an exploratory experiment searching for emergent platform properties. *IEEE Transactions on Software Engineering*, 37(1), 95–108. <https://doi.org/10.1109/TSE.2010.22>
- Purao, S., Bolloju, N., & Tan, C.-H. (2018). A modeling language for conceptual design of systems integration solutions. *ACM Transactions on Management Information Systems*, 9(2), 1–25. <https://doi.org/10.1145/3185046>
- Puschmann, T., & Alt, R. (2016). Sharing economy. *Business & Information Systems Engineering*, 58(1), 93–99. <https://doi.org/10.1007/s12599-015-0420-2>
- Sakata, I., Sasaki, H., Akiyama, M., Sawatani, Y., Shibata, N., & Kajikawa, Y. (2013). Bibliometric analysis of service innovation research: Identifying knowledge domain and global network of knowledge. *Technological Forecasting and Social Change*, 80(6), 1085–1093. <https://doi.org/10.1016/j.techfore.2012.03.009>
- Schreieck, M., Wiesche, M., & Krcmar, H. (2016). Design and governance of platform ecosystems - key concepts and issues for future research. *24th European Conference on Information Systems (ECIS)*.
- Schwienbacher, A., & Larraalde, B. (2010). Crowdfunding of small entrepreneurial ventures. *SSRN Electronic Journal*. Advance online publication. <https://doi.org/10.2139/ssrn.1699183>
- Sedera, D., Lokuge, S., Grover, V., Sarker, S., & Sarker, S. (2016). Innovating with enterprise systems and digital platforms: A contingent resource-based theory view. *Information & Management*, 53(3), 366–379. <https://doi.org/10.1016/j.im.2016.01.001>
- Sesay, A., Oh, O.-O., & Ramirez, R. (2016). Understanding Sociomateriality through the Lens of assemblage theory: Examples from police body-worn cameras. *37th International Conference on Information Systems (ICIS)*.
- Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482. https://doi.org/10.1007/978-3-642-27922-5_23
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). MIT press. <https://doi.org/10.7551/mitpress/12107.001.0001>
- Simpson, T. W. (2004). Product platform design and customization: Status and promise. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 18(1), 3–20. <https://doi.org/10.1017/S0890060404040028>
- Sørensen, C., de Reuver, M., & Basole, R. C. (2015). Mobile platforms and ecosystems. *Journal of Information Technology*, 30(3), 195–197. <https://doi.org/10.1057/jit.2015.22>
- Spagnoletti, P., Resca, A., & Lee, G. (2015). A design theory for digital platforms supporting online communities: A multiple case study. *Journal of Information Technology*, 30(4), 364–380. <https://doi.org/10.1057/jit.2014.37>
- Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology & Human Values*, 35(5), 601–617. <https://doi.org/10.1177/0162243910377624>
- Star, S. L., & Griesemer, J. (1989). Institutional ecology, ‘translations’ and boundary objects: Amateurs and professionals in Berkeley’s Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science*, 19(3), 387–420. <https://doi.org/10.1177/030631289019003001>
- Stummer, C., Kundisch, D., & Decker, R. (2018). Platform launch strategies. *Business & Information Systems Engineering*, 60(2), 167–173. <https://doi.org/10.1007/s12599-018-0520-x>
- Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of modeling in data mining: Interpretive approach to bankruptcy prediction. *Journal of Management Information Systems*, 16(1), 63–85. <https://doi.org/10.1080/07421222.1999.11518234>
- Sutherland, W., & Jarrahi, M. H. (2018). The sharing economy and digital platforms: A review and research agenda. *International Journal of Information Management*, 43, 328–341. <https://doi.org/10.1016/j.ijinfomgt.2018.07.004>
- Syed, S., & Weber, C. T. (2018). Using machine learning to uncover latent research topics in fishery models. *Reviews in Fisheries Science & Aquaculture*, 26(3), 319–336. <https://doi.org/10.1080/23308249.2017.1416331>
- Tarafdar, M., & Davison, R. M. (2018). Research in information systems: Intra-disciplinary and inter-disciplinary approaches. *Journal of the Association for Information Systems*, 19(6), 2. <https://doi.org/10.17705/1jais.00500>
- Taudes, A., Feurstein, M., & Mild, A. (2000). Options analysis of software platform decisions: A case study. *MIS Quarterly*, 24(2), 227. <https://doi.org/10.2307/3250937>
- Thomas, L., Autio, E. and Gann, D. (2014). Architectural leverage: putting platforms in context. *The Academy of Management Perspectives*, 28(2): 198–219. <https://doi.org/10.5465/amp.2011.0105>
- Thomas, N. T., & Azhuvath, R. M. (2018). Automatic visualization of product features using LDA and Word2Vec. *Proceedings of the International Conference on Computing and Communication Systems 2016*, India.
- Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Research commentary—Digital infrastructures: The missing IS research agenda.

- Information Systems Research*, 21(4), 748–759. <https://doi.org/10.1287/isre.1100.0318>
- Tilson, D., Lyytinen, K. and Sørensen, C. (2010b). Digital infrastructures: the missing IS research agenda. *Information Systems Research*, 21(5):748–759. <https://doi.org/10.1287/isre.1100.0318>.
- Tiwana, A., Konsynski, B., & Bush, A. (2010). Research commentary—platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research*, 21(4), 675–687. <https://doi.org/10.1287/isre.1100.0323>
- Urquhart, C. (2013). *Grounded theory for qualitative research: A practical guide*. SAGE Publications, Ltd. <https://doi.org/10.4135/9781526402196>
- Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, 68(1), 1–17. <https://doi.org/10.4324/9781315699035-9>
- Vargo, S. L., & Lusch, R. F. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, 44, 5–23. <https://doi.org/10.1007/s11747-015-0456-3>
- Vijayarani, S., Ilamathi, J., & Nithya, S. (2015). Preprocessing techniques for text mining—an overview. *International Journal of Computer Science & Communication Networks*, 5(1), 7–16.
- vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (2015). Standing on the shoulders of giants: Challenges and recommendations of literature search in information systems research. *Communications of the Association for Information Systems*, 37. <https://doi.org/10.17705/1CAIS.03709>
- Wade, J. T., Roth, P. L., Thatcher, J. B., & Dinger, M. (2020). Social media and selection: Political issue similarity, liking, and the moderating effect of social media platform. *MIS Quarterly*, 44(3), 1301–1357. <https://doi.org/10.25300/MISQ/2020/14119>
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly*, 26(2), xiii–xxiii. <https://doi.org/10.2307/4132319>
- Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). Research commentary—the new organizing logic of digital innovation: An agenda for information systems research. *Information Systems Research*, 21(4), 724–735. <https://doi.org/10.1287/isre.1100.0322>
- Zhu, K., Dong, S., Xu, S. X., & Kraemer, K. L. (2006). Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of European companies. *European Journal of Information Systems*, 15(6), 601–616. <https://doi.org/10.1057/palgrave.ejis.3000650>

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