Rising like a Phoenix: Emerging from the Pandemic and Reshaping Human Endeavors with Digital Technologies ICIS 2023

General IS Topics

Dec 11th, 12:00 AM

Mapping the Metaverse – Knowledge Generation Structures in a Nascent Ecosystem

Henrik Pohsner
University of Kassel, henrik.pohsner@uni-kassel.de

André Hanelt
University of Kassel, hanelt@uni-kassel.de

Mapping the Metaverse – Knowledge Generation Structures in a Nascent Ecosystem

Completed Research Paper

Henrik Pohsner
University of Kassel
Digital Transformation Management
Henschelstraße 2, 34109 Kassel
henrik.pohsner@uni-kassel.de

André Hanelt
University of Kassel
Digital Transformation Management
Henschelstraße 2, 34109 Kassel
hanelt@uni-kassel.de

Abstract

The so called Metaverse is among the most popular emerging themes in contemporary business and academic discourses related to digital transformation. Beside the hype and buzz around the topic, recent research has established important progress related to the technical foundations and use-case scenarios. However, the preconditions to manifest these promising potentials, that is the underlying socio-technical ecosystem, is, to date, not well-understood. In particular, we still do not know much about the key building blocks of technical knowledge and who creates this knowledge base necessary for the Metaverse. Therefore, in this study, employing cluster-analysis to a complex dataset involving 2297 Metaverse-patents, we reveal key areas of technological knowledge as well as important actor groups shaping the development of the knowledge base. Building upon these inductively derived knowledge generation structures, we discuss important implications for managers and academics and propose an Information Systems research agenda to further explore the emerging Metaverse ecosystem.

Keywords: Metaverse, Nascent Ecosystems, Knowledge Generation, Patent Analysis, Cluster Analysis

Introduction

Since September 2021, Google Trends for the search term “Metaverse” skyrocketed from nearly zero up to 100 in January 2022 indicating the topic’s popularity. Companies have also started to evaluate how Metaverse applications can create business value, by either offering a new revenue source or optimizing processes: For instance, BMW is said to have built a virtual factory in the Metaverse (Diaz 2023). Another example is Fortnite, a computer game by the developer firm Epic Games, which offers a variety of virtual experiences in the Metaverse (Seidel et al. 2022). However, there are also examples that indicate uncertainty regarding the future of the Metaverse: Meghan Bobrowsky, a journalist of The Wall Street Journal even called it the “Meh-taverse”, as companies like Microsoft and Meta started to streamline their Metaverse related departments in their latest cost cutting initiatives (Bobrowsky 2023). Walt Disney even closed their Metaverse department, which was planned to develop strategies and content for upcoming Metaverse applications (Whelan and Flint 2023).

Aside from the hype and buzz we witness in business media these days, there is still a lack of clarity as to what the Metaverse actually is, how it can be defined (Park and Kim 2022; Peukert et al. 2022) and whether the Metaverse already exists (Schöbel and Leimeister 2023). Accordingly, academic work has picked up the topic, commented and projected its potentials as well as related challenges (e.g. Dincelli and Yayla (2022) and Schöbel and Leimeister (2023)). For instance, Dincelli and Yayla (2022) identified applications of the Metaverse in literature and investigated immersive virtual reality (VR) and its opportunities and challenges.

Schöbel and Leimeister (2023) conceptualized the Metaverse as a meta-ecosystem and derived opportunities for further research in this context. Albeit these valuable contributions, highlighting particular aspects and potentials, what is missing to date is a systematic assessment of the existing substance of the Metaverse. One way to do this is an analysis of the established knowledge base underlying the Metaverse. What we do know from current work is that Metaverse applications build upon a complex assemblage of technologies and associated heterogeneous actors (Dincelli and Yayla 2022; Dolata and Schwabe 2023). Therefore, scholars unanimously point to the value of the ecosystem lens for understanding the phenomenon (Schöbel and Leimeister 2023; Seidel et al. 2022). Indeed, research about digital innovation ecosystems (Wang 2021) may lend some insights and ideas for structuring the Metaverse and capturing its dynamics. At the same time, the Metaverse and its current state offer a unique opportunity to learn about the emergence of modern digital innovation ecosystems by focusing the underlying knowledge base. Prior Information System (IS) research has generated rich insights about mature ecosystems such as enterprise software and the role of knowledge for their success (Ceccagnoli et al. 2012; Haki et al. 2022; Huang et al. 2009).What we do know is that, for nascent ecosystems, the cultivation of the underlying knowledge base is crucial for its further development (Moeen et al. 2020; Moeen and Agarwal 2017; Moore 1993): This may hold true even more for digital innovation ecosystems as, while for all innovations the integration of heterogeneous knowledge is crucial, with digital innovation, the heterogeneity of knowledge and accordingly the demand for balancing and integrating knowledge (Yoo et al. 2012) increases.

In this study, we investigate the following research question: *How is the knowledge base of the emerging Metaverse ecosystem constituted?* Finding answers to this question is important for both academic and practical reasons. On the one hand, the current academic discourse may profit from more systematic empirical analysis to move beyond the state of generalized commentaries, visions or niche investigations. Apart from the context of the Metaverse, IS research needs to advance its understanding of nascent ecosystems that build upon digital technologies (Wang 2021). On the other hand, managers may better navigate strategic questions pertaining to the Metaverse on the basis of a systematic assessment of existing key technology areas as well as key actors shaping the knowledge base. Accordingly, we collected a comprehensive dataset on Metaverse knowledge by deriving patent data from the United States Patent and Trademark Office (USPTO) patent database. The dataset was enriched with the Standard Industrial Classification (SIC) code for each firm holding a patent to determine from which industry sector the patent holder originally comes from. Based on this dataset we conducted two step-cluster analyses to identify technology clusters as well as clusters of firms creating knowledge in the identified technology clusters. We identified six technology clusters, which are focused either on a specific application or on technological infrastructures for the Metaverse and seven clusters of firms contributing knowledge to one of these technology clusters or, in the case of one particular cluster, to all technology clusters.

Our findings contribute to theory by providing a first exploratory, yet systematic examination of the knowledge base that makes up the Metaverse including the associated actors. To advance IS knowledge related to the phenomenon, we discuss existing, often conceptual work, in the light of our empirical findings. Furthermore, we contribute to digital innovation ecosystem literature by revealing knowledge generation structures as an important means to understand nascent instances (Moeen et al. 2020). Additionally, we underscore and extend existing work on the digital innovation ecosystem holarchy (Wang 2021). Moreover, our results offer valuable insights for practitioners as we identified an existing knowledge base for Metaverse infrastructure and applications, created by heterogeneous actors. Firms can use our findings to pro-actively check possible strategies and evaluate various diversification opportunities offered by the Metaverse. Our findings further allow firms to evaluate which knowledge already exists, assess which further opportunities for knowledge generation remain and identify possible partners with existing knowledge.

Theoretical Foundations

The Metaverse in the Academic and Practical Discourse

The term Metaverse was coined by Neal Stephenson in his science fiction novel Snow Crash in 1992. In the dystopian world of the novel, the Metaverse is a virtual world, accessible with specific hardware (Dolata and Schwabe 2023; Seidel et al. 2022; Stephenson 1992). Apart from the novel, the Metaverse-concept also receives attention in scientific literature. Park and Kim (2022) described the Metaverse as a "threedimensional virtual world where avatars engage in political, economic, social, and cultural activities" (p.4211). This definition was recently advanced by Schöbel and Leimeister (2023) who described the Metaverse as "a massively scaled and interoperable meta ecosystem of other digital ecosystems of real-time rendered 3D virtual worlds that can be experienced synchronously and persistently by an unlimited number of complementors and consumers with an increased user experience caused by a creativity-guided co-creation of goods managed by orchestrators and supported by platform owners" (Schöbel and Leimeister 2023, p.7). In literature, the Metaverse has been investigated from different perspectives, for example Dolata and Schwabe (2023) identified three main perspectives in research, which are namely technical, use-oriented, and structural. In what follows we structure existing literature according to these perspectives.

The development of the Metaverse from a technological perspective is divided into three stages by Dolata & Schwabe (2023): In a first stage, virtual worlds were hosted on a computer and ended as soon as the computer is turned off. The second stage involved well known multiplayer online games, made possible by faster internet connections and progress in 3D graphics. The third stage begun with software and hardware advancements and developments since the late 2000's. One crucial aspect for the implementation of the Metaverse are therefore different hardware and software technologies (Dincelli and Yayla 2022; Peukert et al. 2022), especially sufficient hardware power and engines for 3D virtualizations (Dolata and Schwabe 2023). Those technologies are the basis for features of a successful Metaverse implementation such as immersive realism, ubiquity of access and identity, interoperability, and scalability (Dionisio et al. 2013). Dionisio et al. (2013) described immersive realism as the realism of the virtual world that makes users feel immersed in it. Ubiquity means the accessibility with several devices like PCs and mobile devices, and the assurance that the virtual identity will be preserved as the user transitions in the Metaverse. Interoperability describes standards for digital assets and the possibility to move between locations, whereas scalability describes an architecture with enough power for a Metaverse with a huge number of users. These features of a successful Metaverse highlight, that there are several technologies needed to build and operate a Metaverse. One type of required technology can be summarized as extended reality (XR) (Mystakidis 2022). According to Mystakidis (2022) this includes Virtual Reality (VR), Augmented Reality (AR) and Mixed Reality (MR). These technologies allow to access and experience the Metaverse and are the foundation of immersive virtual reality applications (Dincelli and Yayla 2022). Other technologies that ensure ownership of virtual goods and virtual properties and enable virtual transactions between users of the Metaverse are required, which is possible with nonfungible tokens (NFTs) (Dolata and Schwabe 2023; Mystakidis 2022).

The use-oriented perspective is focused on the user and field of applications. Gaming is a frequently named field of application and seen as "the most common platform in the popularization of the Metaverse" (Park and Kim 2022, p.4225). Interestingly, gaming providers like Epic Games also offer experiences in their games different to the original purpose of the game, e.g. concerts in Fortnite in separated environments (Seidel et al. 2022). Besides gaming (e.g. Park and Kim (2022)), different office applications (Xi et al. 2022), entertainment, travel and retail (Dincelli and Yayla 2022), social interaction, marketing and education (Park and Kim 2022) or healthcare (Zahedi et al. 2022) are among the often bespoken fields of application. Activities previously located in the physical world could be completely handled in the virtual world by using XR devices, e.g. an appointment at the real estate agent followed by driving to work afterwards can be replaced by the Metaverse as both can be virtually conducted by using XR technology in the office (Schöbel and Leimeister 2023).

Finally, a stream in the literature approaches the Metaverse from a structural perspective. Due to the aforementioned variety in technologies as well as heterogeneity in actors that need to engage in a complex interplay, research in this areas has strongly emphasized the value of an ecosystem perspective (Schöbel and Leimeister 2023; Seidel et al. 2022). The ecosystem perspective on the Metaverse is valuable because several actors from different industries and ecosystems are involved in the creation and operation of the Metaverse. As described above, heterogeneous actors must interact as, for instance, hardware is delivered by other actors than virtualization engines and experiences by other actors as well (Dolata and Schwabe 2023). Furthermore, Seidel et al. (2022) outlined that every unique universe the Metaverse is composed of has its own infrastructure that allows it to keep and convert digital assets like skins or currencies, also when transitioning between different universes. Schöbel and Leimeister (2023) further highlighted that the Metaverse as a meta-ecosystem connects several physical and virtual worlds as an intersection between them. Similarly, scholars perceive the Metaverse as a network consisting of multiple virtual worlds (Dionisio et al. 2013) or further an ecosystem connecting several ecosystems (Schöbel and Leimeister 2023; Seidel et al. 2022). The Metaverse is described as not being one single virtual world, but as a network of virtual worlds which offer several different experiences (Dionisio et al. 2013; Seidel et al. 2022) and is accordingly viewed as a construct which could become an *"ecosystem of ecosystems"* (Dolata and Schwabe 2023, p.36). Therefore, the Metaverse not only seems to be suitable to apply existing knowledge on digital innovation ecosystems, but also offers the possibility to learn about the interaction of vertically and horizontally dispersed ecosystems (Wang 2021). However, while different technologies were identified, their relevance was highlighted and potential actors were named, a systematic analysis of how the ecosystem could be structured is missing. Building upon the current state of the literature, it seems that the technology is available, and possible user-applications were identified, but there seems to be less knowledge on the structural foundations, which may result in a bottleneck for the realization of the Metaverse.

Nascent Digital Innovation Ecosystems

An ecosystem in a business context is a metaphor derived from biology and describes networks consisting of several companies interacting cooperatively and competitively (Moore 1993). Ecosystems are characterized, inter alia, by heterogeneous actors (Yoo et al. 2010), which share the goal of joint value creation (Adner 2017). Innovation ecosystems are described as focused on an innovation or value proposition (Jacobides et al. 2018). Digital innovation, that is, new products, services and processes based on digital technologies (Nambisan 2017), have further emphasized the importance of innovation ecosystems due to their interconnected, distributed, and combinatorial nature (Yoo et al. 2010, 2012). As to these traits, digital technologies and the actors creating and participating in digital innovation stem from different industry backgrounds and can be found on different architectural layers (Selander et al. 2010; Wang 2021; Yoo et al. 2010). Managing them is a challenge for organizations (Baldwin 2012), which is why firms are changing the way of how they organize and why ecosystems could be an appropriate form (Wang 2021) to leverage the innovation strategies of heterogeneous actors (Yoo et al. 2010). Wang (2021) defined digital innovation ecosystems as "*a loosely coupled set of autonomous actors (people and organizations who interact without hierarchical fiat) involved in the development and implementation of innovations enabled by digital technologies"* (Wang 2021, p.397), highlighting two main constituents of digital innovation ecosystems, namely the actors and technology (Nischak et al. 2017).

Importantly, the high level of interconnectedness involved in digital innovation puts a premium on extending the perspective even more to include not only a particular ecosystem but also its connections to vertically or horizontally dispersed associated ecosystems (Davis 2016; Wang 2021). While this next step of digital innovation ecosystem research is still in its infancy, first empirical work has illustrated the value of an inter-ecosystem perspective (Wang 2021). For instance, Biedebach et al. (2021) find evidence for the influence of the IT ecosystem on the automotive ecosystem. Furthermore, actors can also benefit from a diversified strategy by not investing only in one ecosystem, but splitting the operations across multiple ecosystems (Selander et al. 2013). Besides, also ecosystems can be a part of other ecosystems. Wang (2021) developed a "holarchy of digital innovation ecosystems" (Wang 2021, p.402), showing entities (holons) as parts of ecosystems on different levels of the holarchy with Product/Service, Business/Entrepreneurial and Category Ecosystems. Viewing digital innovation ecosystems from this perspective highlights the multilevel nature (Wang 2021): First, digital technologies facilitate innovation on different levels of the digital innovation ecosystem and second, ecosystems are at the same time part of a higher-level ecosystem (part) and a higher-level ecosystem for underlying ecosystems (whole).

The development of an ecosystem is a process following multiple steps. According to Moore (1993), ecosystems in a business context undergo several distinct stages, namely Birth, Expansion, Leadership and Self-Renewal, each of which is characterized by specific challenges. Of these stages, Birth and Expansion are the ones pertaining to nascent ecosystems, as both are related to steps necessary for building and scaling an ecosystem, while Leadership and Self-Renewal are related to further improving an existing ecosystem. A similar distinction was also made by Han et al. (2022) who further noted a limited co-evolution between actors and low monetization in the Birth stage and fast growth in the Expansion stage, while the ecosystem still can fail. While substantial research in existing IS literature has been devoted to already existing and mature ecosystems such as enterprise software (Ceccagnoli et al. 2012; Huang et al. 2009, 2018), the literature for new and emerging digital innovation ecosystems is scarce. A notable exception is Hodapp et al. (2019), who investigated nascent digital platform ecosystems and their value co-creation challenges, as value creation differs between mature and nascent ecosystems due to specific characteristics (Hodapp et al. 2019): In nascent ecosystems unexpected innovation can lead to technological challenges, it is unclear who will be the competitors and uncertainty regarding the availability of required resources (Hodapp et al. 2019). Further, actors from different ecosystems can have difficulties to agree on the design of a nascent ecosystem (Du 2018; Hodapp et al. 2019). In nascent ecosystems a primary challenges for actors are the definition of a value proposition and the protection of an idea to prevent other firms from creating a similar offer (Moore 1993). This observation highlights the value of a knowledge-based perspective on nascent ecosystems. We know from existing research on established ecosystems that the underlying knowledge base is important for the development of an ecosystem as a whole (e.g. Huang et al. (2018)) and that digital innovations are based on heterogeneous knowledge bases (Yoo et al. 2012). However, while mature ecosystems build upon knowledge recombination, knowledge creation is particularly relevant for nascent ecosystems. Also aside from IS literature authors perceived these differences of nascent and mature stages of ecosystems (Adner and Kapoor 2010; Dattée et al. 2018; Han et al. 2022; Jacobides et al. 2018): Dattée et al. (2018) noted that the sharing of existing intellectual property may be necessary in order to develop an ecosystem to the desired direction. In contrast, in nascent stages of ecosystems, firms are dependent on knowledge creation of other actors, as “suppliers” must develop required components for a solution before it can be brought to market (Adner and Kapoor 2010). Accordingly, authors noted that ecosystems can emerge from knowledge ecosystems if the innovation from those can be commercialized (Clarysse et al. 2014). Albeit these potentials, we do not know much about the knowledge creation in emerging digital innovation ecosystems. The emerging Metaverse presents an opportunity to observe and analyze a modern emerging digital innovation ecosystem and the underlying knowledge base to advance both digital innovation ecosystem research and the scholarly discourse on the phenomenon.

Methodology

Data Collection

Our analysis of knowledge building for the Metaverse is based on patent data. Patent data is a widely used data source for analyzing structures in knowledge generation (Ardito et al. 2018; Lou and Wu 2021; Stephan et al. 2017). Past empirical work has shown the importance of patents for digital innovation (Hanelt et al. 2021). Furthermore, Moore (1993) mentioned in the context of emerging ecosystems that protection of an idea to prevent other firms to create a similar offer is an important challenge at a particular stage of ecosystem development. Patents as a form of intellectual property are a suitable way to protect knowledge, used especially in the incubation stage of emerging industries, as they can be used to prevent the development of similar technologies (Moeen et al. 2020). Additionally, other authors investigating knowledge structures also relied on patent data (Bierly and Chakrabarti 1996).

We derived the patent data from the patent search of the USPTO, which provides access to patent data published by the USPTO (Chiang et al. 2017). We specifically used the USPTO advanced search, which includes all US patents but also patents from other patent offices around the world, if they are also registered in the USA. Forman and van Zeebroeck (2011) used the USPTO database to enrich their dataset of business IT investment data with patent data. The USPTO participates in the electronic priority document exchange (PDX) program, meaning that several patent office’s exchange priority documents electronically (USPTO 2023a). In practice it means a lot of patents are registered worldwide, as companies fill patent applications for several patent offices. For example, Beijing Baidu Netcom Science and Technology Co Ltd as a Chinese company registered a patent at the USPTO, but also in Japan and Korea. The USPTO is also part of the Patent Cooperation Treaty (PCT), which allows the filing of a patent in several countries by filling only one international patent application (USPTO 2023b). We searched the USPTO database for the term “Metaverse”, assuming that a relevant patent mentions the Metaverse, and derived both, already granted and pre-granted patents. Other researchers also applied keyword-based searches to extract patents (Battke et al. 2016; Costantini et al. 2015), but added patent classification codes or created a list with appropriate keywords. We decided to use only the keyword, because we aimed to search exploratively for patents related to the Metaverse and this approach allows a clear focus. We further assume that the Metaverse cannot be limited to specific classifications and a list with keywords would not be exhaustive. To enhance the database, we added cited patents as these are technologies related to the found patents and therefore also relevant for the knowledge base, which resulted in a dataset of 2297 patents. As patent application procedures can be tedious, it makes sense to also include those patent applications which are currently evaluated to ensure that also relatively new knowledge is included in the dataset. Regardless of whether the patent is a granted or a pre-granted patent, all patent records contain the title of the patent, an abstract, a detailed description, applicant information, assignee information and the IPC/CPC codes related to the patent. Additionally, they also have a document ID which relates to the patent ID or the ID of a pre-granted patent and an application ID which is unique and referenced for granted and pre granted patents. We favored the Cooperative Patent Classification (CPC) codes instead of International Patent Classification (IPC) codes as CPC is an extension of the IPC code and jointly managed by the European Patent Office and USPTO. Based on the application ID we ensured that each patent is listed only one time in the dataset. We removed accordingly duplicates and all pre-granted patents if the granted patent is already included in the dataset. Furthermore, we determined the patent holder by checking if it is currently assigned. If yes, we choose the assignee. If not, we checked if there is an applicant we can choose as patent holder. For some patents, there is no assignee or applicant because only the inventor, i.e., a person and not a company is assigned. In those cases, we removed the patent from the dataset. In a further step we checked the entity holding the patent. Occasionally, patent holders are the same but written differently. For example, there is the notation “Corp”, “Corp.” and “Corporation”, all of which have the same meaning, so we have chosen one notation for each case to ensure that the firm is consistently named. In a next step, for each company holding a patent, the four-digit SIC Code was determined and assigned. SIC codes are a standard in industry classification which is used to determine the sector in which a patentholder is active (Stephan et al. 2017). We used the primary SIC code, because we aimed to identify from which industries the patentholders are originally from and not all activities of all business units. However, it was not possible to add a SIC code for every company, as several companies have no publicly available descriptions of their business or only an entry in a company register which does not include the primary sector of the company. For those companies no SIC code was assigned.

Data Analysis

The purpose of this study is to investigate the knowledge generation structures of the Metaverse by identifying the knowledge base from a technology perspective and the firms developing this knowledge. Our dataset offers several opportunities to investigate the technological knowledge and the actors creating or holding this knowledge. An appropriate approach for identifying technology and actor groups is cluster analysis, as applied by other authors who used, inter alia, patent data in their studies (Bierly and Chakrabarti 1996; Wacker et al. 2014). We decided to apply two-step cluster analyses, similar to studies seeking to segment groups (Rundle-Thiele et al. 2015; Tkaczynski 2017), using the software SPSS to identify patterns in the dataset regarding the technology (CPC codes) and firms. Both, cluster analysis and SPSS are widely used by researchers to identify patterns in datasets and cluster homogeneous groups (e.g. Biedebach and Hanelt (2020)). We applied a two-step clustering for both cluster analyses as it is suitable for large data sets (e.g. Rundle-Thiele et al. (2015)) and able to find groups in datasets with a similar quality as other classification methods (Kent et al. 2014). Additionally, it is suitable for explorative analysis (Tkaczynski 2017) as needed for this study: Two-step clustering includes automatic determination of the most suitable number of clusters (IBM 2022; Tkaczynski 2017), which is an advantage over K-Means clustering for this particular dataset. These clusters might be not a perfect result if considering the context, but at least a very good basis. Further the two-step clustering assumes that variables of the cluster analysis are independent. The variables we use to cluster our data points (patents) based on their similarities, are each one CPC code, which is either assigned to the patent or not. We therefore assume the independence of variables is given. The same is true for the second cluster analysis which clusters the firms based on their involvement in the resulting clusters. In general this method follows two distinct steps (Rundle-Thiele et al. 2015): First, a cluster features tree is constructed (Okazaki 2007; Rundle-Thiele et al. 2015) to create preclusters, which are, in a second step, clustered by a hierarchical clustering algorithm that allows to evaluate solutions with different numbers of clusters (Norušis 2011). In the model summary view for two step clustering the quality of a cluster can be evaluated based on the silhouette measure of cohesion and separation (IBM 2021): It indicates whether the results are “poor” (no significant evidence), “fair” (weak evidence) or good (reasonable or strong evidence) based on Kaufman and Rousseeuw (1990) rating for clusters. This measure allows us to decide whether resulting clusters are sufficiently distinguishable and simultaneously how similar the included objects of a cluster are. The silhouette measure of cohesion and separation measures averages over the complete dataset:

$$\frac{B - A}{\max(A, B)}$$

A is the distance to the cluster center of the cluster the specific object is assigned to, and B is the distance to the center of the nearest cluster it is not assigned to. The results can range from -1.0 (all values are on the cluster center of another cluster) over 0.0 (equidistant to own and nearest cluster center) and 1.0 (all directly on own cluster center). This means, the higher the coefficient is, the better objects of a cluster match to its assigned cluster while simultaneously not match the nearest cluster, which indicates a higher cluster quality. However, this does not mean that only clusters with a silhouette coefficient with 1.0 are good. Lower values also indicate that objects of a cluster fit to their assigned cluster, even though not perfectly. The first cluster analysis aimed to identify technology clusters in the context of the Metaverse. In order to identify these technology clusters, we performed a cluster analysis for CPC codes assigned to each patent. These CPC codes as a kind of technology code allow us to identify technology-based clusters. While the two-step cluster method does not need a transformation of data for statistical reasons (Rundle-Thiele et al. 2015), we needed to reduce the complexity of CPC codes. CPC codes are very detailed as each code is for a specific technology and application, but they are also clearly structured in sections, classes, subclasses and groups or subgroups. To reduce the complexity, we used the subclass instead of the complete CPC code. The subclass is already quite detailed and can contain hundreds of specific complete CPC codes. Additionally, we do not aim to identify clusters with very specific technology application, but clusters with similar technology categories. Accordingly, we first cut off the digits after the first four digits and second, created one variable for each unique CPC subclass. For the two-step cluster analysis, we coded these variables binary with "0" for not assigned and "1" for assigned to the patent. The two-step cluster analysis uses the log-likelihood distance as a distance measure for categorical variables determining the similarity between clusters, assuming that the categorical variables are multinomial distributed (IBM 2022). Two-step clustering offers the possibility to automatically determine the best number of clusters for the dataset (Rundle-Thiele et al. 2015). However, this is solely based on statistical measures like the Bayesian Information Criterion or the Akaike Information Criterion and not on further considerations taking the context into account (IBM 2022). We first performed cluster analysis with a number of clusters determined in the first step of the two-step cluster analysis based on either the Bayesian Information Criterion or the Akaike Information Criterion and then evaluated the results. The statistical quality of the resulting clusters was defined as "fair", indicating acceptable results (Rundle-Thiele et al. 2015). However, the results must be also evaluated in the context of the data and aim of analysis. Cluster analysis as an exploratory approach does not necessarily lead to one perfect result, it is also possible that there are more adequate solutions (Biedebach and Hanelt 2020). Therefore, we decided to test further cluster analysis with different number of clusters that have similar cluster qualities compared to the results of the automatically determined number of clusters. We take existing Metaverse literature into account when evaluating the results, in which technologies, applications and actors where identified (Dincelli and Yayla 2022; Dolata and Schwabe 2023). After carefully evaluating the results, we identified six clusters as the best solution from the statistical and contextual perspective. The silhouette measure of cohesion and separation of this cluster analysis is 0.3, which is defined as a "fair" quality of the clusters meaning the indicated cluster quality is in the mid-range and a sufficient result in terms of heterogeneity between clusters and homogeneity within the cluster: A silhouette coefficient of 0.3 indicates that the data points are on average not all directly on the center of their clusters and not equidistant between the own and nearest other cluster center. That is, the resulting clusters are not completely distinct, but there are enough differences between clusters and similarities within clusters to speak of distinguishable groups. The silhouette measure of cohesion and separation need to be above a level of 0.0 for a sufficient distinction between clusters (Rundle-Thiele et al. 2015).

The second cluster analysis was conducted with the aim of identifying companies that are similar in terms of the patents they hold in the technology clusters identified in the first cluster analysis. For the analysis, we added the results of the first cluster analysis to the original data set: First, we created a list of all the companies in the dataset and second, we added the clusters of the first cluster analysis as new characteristics of the firms in the dataset. We then summed up the number of patents each firm holds in each of the technology clusters and added these sums to the dataset. For the second cluster analysis we again applied the two-step clustering. Each cluster of the previous cluster analysis is one variable for this cluster analysis, containing the number of patents a firm holds in this cluster, which are continuous data. For this type of data, we again used the log-likelihood as distant measure, but this time for continuous data. We did not specify a minimum or maximum number of clusters for the analysis. Further, we needed to choose an information criterion for the determination of the number of clusters. We decided to choose the Akaike Information Criterion as it is a widely used information criterion in model selection and suitable for large datasets (Cavanaugh and Neath 2019). We again evaluated the resulting clusters regarding their contextual quality. This time the cluster quality was good, both statistically and contextually. Therefore, we decided not to choose a number of clusters that deviates from the automatically determined number. The silhouette measure of cohesion and separation is 0.7 for this cluster analysis, which is defined as a "good" quality of the clusters. A silhouette coefficient of 0.7 indicates that the data points are on average near on the center of their clusters. That is, the resulting clusters are quite distinct and sufficient differences between clusters and similarities within clusters exist to speak of distinguishable groups.

Results

In this section we describe the findings of both cluster analyses, based on the CPC codes in technology clusters and SIC codes in the firm clusters. The first two-step cluster analysis for the CPC subclasses revealed six technology clusters: For more precise technology information we decided to use the more specific CPC groups to describe the resulting clusters. CPC groups are included if they are above a threshold of 15% of all CPC codes assigned to a patent within the cluster to ensure a sufficient relevance of the CPC group. It should be noted that there is a wide variety of further CPC groups which we do not mention here. In the following we give a brief description of the clusters:

Cluster	Description
Technology Cluster 1: "Data Infrastructure"	As the main CPC groups are related to applications like information retrieval, databases and transferring data, we decided to name this cluster "Data Infrastructure". The CPC codes suggest that it is an infrastructure cluster with patents related to hardware and software inventions primarily used in the context of data and databases.
Technology Cluster 2: "Digital Entertainment"	As the main CPC groups are related to interactive television, video on demand, data transferring, 3D image rendering, image analysis and 3D models and images for computer graphics, we decided to name this cluster "Digital Entertainment". The CPC codes suggest that it is an application cluster with patents related to technologies used for digital entertainment offerings.
Technology Cluster 3: "Network and Security Infrastructure"	As the main CPC groups are related to cryptographic arrangements for secure communication and network services, network security and other security arrangements, we decided to name the cluster "Network and Security Infrastructure". The CPC codes suggest that it is an infrastructure cluster with patents related to technologies used for network and data security.
Technology Cluster 4: "Healthcare"	As the main CPC groups are related to applications like diagnostic, devices and apparatus bringing media into the body and sensors and other devices, we decided to name this cluster "Healthcare". The CPC codes suggest that it is an application cluster with patents related to technologies used for healthcare related technology.
Technology Cluster 5: "Gaming"	As the main CPC groups are related to video games and input arrangements transferring data in a form the computer can handle, as well as network services are represented in this cluster, we decided to name this cluster "Gaming". The CPC codes suggest that it is an application cluster with patents related to technologies for hardware and software used for gaming applications.
Technology Cluster 6: "E-Commerce and Asset Trading".	As the main CPC groups are related to commerce, payment architecture and finance, we decided to name this cluster "E-Commerce and Asset Trading". The CPC codes suggest that it is an application cluster with patents related to technologies used for different asset exchange and e-commerce applications.

Table 1. Results of First Cluster Analysis

The second cluster analysis based on the clusters of the first cluster analysis revealed seven firm clusters. To describe which firms can be found in the cluster, we used the SIC codes and evaluated how often a SIC code can be found in the respective cluster. We used the three digit SIC codes as we aggregated SIC codes (Stephan et al. 2017). All clusters are characterized by a diversity of actors, visible through a wide variety of SIC codes. While the dataset includes a lot of firms with computer programming, data processing, and other computer related services SIC codes (737) or firms with no assignable SIC code, it is nevertheless possible to identify a certain trend within the clusters. In the following table the industry focus and exemplary firms can be found:

Cluster	Main industries based on 3-digit SIC codes	Examples
Firm Cluster 1	data processing and communication.	ZTE CORPORATION, JPMorgan Chase Bank, Mastercard International Incorporated and Coinbase, Inc.
Firm Cluster 2	computer programming, data processing, computer equipment and communication	NVIDIA Corporation, Meta, Roblox Corporation and Tencent Technology (Shenzhen) Company Ltd
Firm Cluster 3	computer programming, data processing and communication equipment, management services and advertising	Nintendo Co., Ltd., Adobe Inc., Konami Co., Ltd. And Disney Enterprises, Inc.
Firm Cluster 4	computer programming, data processing, educational institutions and communication equipment and services	Oculus VR, LLC, Sony Ericsson Mobile Communications AB and Warner Bros. Entertainment Inc
Firm Cluster 5	computer programming, data processing, security and commodity exchange, allied services and security brokers	Bank of America Corporation, The NASDAQ OMX Group, Inc. and Bloomberg L.P
Firm Cluster 6	computer programming, data processing, computer equipment, electronic components	Huawei Technologies Co., Ltd., Taiwan Semiconductor Manufacturing Co., Ltd. And Dell USA, L.P
Firm Cluster 7	medical and laboratory instruments and analytical, optical and measuring instruments	Kimberly-Clark Worldwide, FANUC CORPORATION and Leica Microsystems CMS GmbH

Table 2. Results of Second Cluster Analysis

Each firm cluster is related to a technology cluster: Firms within a firm cluster hold primarily patents in one of the technology clusters, so the clusters are homogenous in terms of technology knowledge created by the firms (see figure 1). An exception is the second firm cluster which consists of firms holding patents in several technology clusters. However, this does not mean that the firms in clusters are homogenous if they hold patents in the same technology cluster, rather they are quite heterogeneous in terms of their industries (SIC codes).

Firm Cluster:	Technology Cluster:	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Sums
Cluster 1	Patents total	5	0	172	0	0	7	184
	Patents [%]	3%	0%	93%	0%	0%	4%	
	Patents per firm	0,03	0,00	1,13	0,00	0,00	0,05	1,21
Cluster 2	Patents total	208	341	317	63	148	131	1208
	Patents [%]	17%	28%	26%	5%	12%	11%	
	Patents per firm	1,70	2,80	2,60	0,52	1,21	1,07	9,90
Cluster 3	Patents total	9	30	22	1	132	16	210
	Patents [%]	4%	14%	10%	0%	63%	8%	
	Patents per firm	0,08	0,28	0,20	0,01	1,22	0,15	1,94
Cluster 4	Patents total	0	164	0	0	4	10	178
	Patents [%]	0%	92%	0%	0%	2%	6%	
	Patents per firm	0,00	1,00	0,00	0,00	0,02	0,06	1,09
Cluster 5	Patents total	0	5	20	0	35	169	229
	Patents [%]	0%	2%	9%	0%	15%	74%	
	Patents per firm	0,00	0,04	0,15	0,00	0,26	1,27	1,72
Cluster 6	Patents total	85	0	3	2	0	18	108
	Patents [%]	79%	0%	3%	2%	0%	17%	

	Patents per firm	1,02	0,00	0,04	0,02	0,00	0,22	1,30
Cluster 7	Patents total	0	17	4	159	0	0	180
	Patents [%]	0%	9%	2%	88%	0%	0%	
	Patents per firm	0,00	0,14	0,03	1,27	0,00	0,00	1,44
Sums	Patents in Cluster	307	557	538	225	319	351	

Table 3: Distribution of Patents in Technology Clusters to Firm Clusters

We want to exemplarily outline these differences with two clusters highlighting the differences in knowledge generation (figure 2 and 3). While firms in cluster 1 have a clear focus on patents that can be found in technology cluster "Network and Security Infrastructure", firms in cluster 2 have patents in multiple technology clusters. It indicates that firms in cluster 2 create knowledge for several technology clusters, while firms in the other cluster are more focused in their knowledge creation as they primarily create knowledge for one technology cluster.

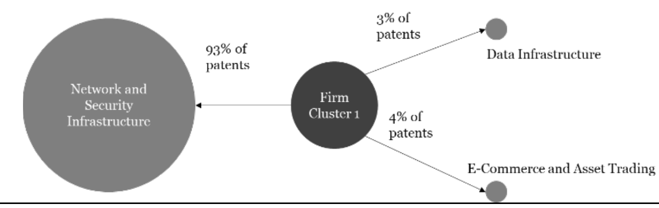

Figure 1: Connection between Firm Cluster 1 and Technology Clusters

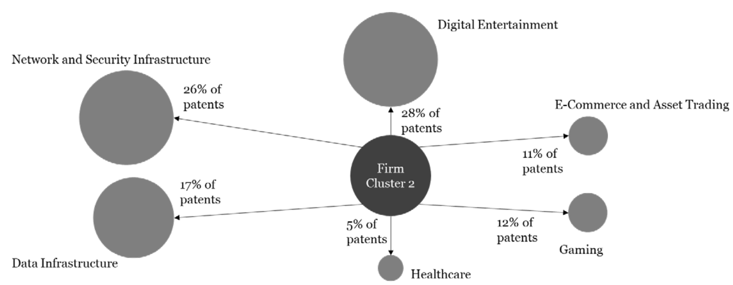

Figure 2: Connection between Firm Cluster 2 and Technology Clusters

Discussion

Technology Perspective

The cluster analysis for CPC subclasses revealed six different technology knowledge areas. These clusters indicate that actors generate knowledge related to several specific fields of application. We identified the application clusters "Gaming", "Digital Entertainment", "Healthcare" and "E-Commerce and Asset Trading". This finding is partly supported by prior Metaverse research that has identified fields of application like gaming, office applications, social applications, marketing, education, travel, entertainment and healthcare (Dincelli and Yayla 2022; Park and Kim 2022; Zahedi et al. 2022). We cannot confirm the existence of all these applications mentioned in literature based on patented knowledge. We assume that this can have at least three reasons: First, it is possible that these applications currently do rather exist in visions or projections but are still away from being realized in practice. Following this train of thought, one could assume that the fields of application we identified reside on a higher maturity level. As an alternative explanation, second, it is possible that no knowledge must be patented for the application, for instance if it is very generic knowledge or the application does not need specific knowledge. Third, it might be that actors do generate knowledge in further fields of application but do not associate their work with the emerging theme of Metaverse.

We found several different technologies in each cluster. It indicates that not only application-specific knowledge, but also general technology knowledge can be found to a certain degree in each cluster. For instance, the cluster "Healthcare" contains a lot of knowledge related to healthcare, but also a few CPC codes indicating knowledge related to network technologies. Dionisio et al. (2013) described immersive realism, ubiquity, interoperability, and scalability as crucial for a successful Metaverse. These characteristics highlight that there are several technologies needed to build and operate the Metaverse, as it requires technology that makes users feel immersed, standards for interoperability and hardware to access and operate the Metaverse. Following the characteristics, we can assume that applications not only demand knowledge specific for the application, but also general technologies for network, visualization, or immersive VR purposes. However, it should be noted that a patent can have multiple CPC codes, which could also lead to patents that are classified as both, very application specific, but also related to network and other technologies. Our study shows that a lot of knowledge related to the Metaverse was and still is created. Therefore we assume it is likely that a large proportion of required components developed by "suppliers" for the Metaverse already exist and uncertainty regarding the technological foundation of the Metaverse is already decreasing (Adner and Kapoor 2010).

Apart from that, we identified two clusters that cannot be associated with a particular field of application. These technology clusters might be best understood as infrastructural clusters and comprise "Data Infrastructure" and "Network and Security Infrastructure". They cover mainly technology related to databases, data transfer and security of connections. In Metaverse literature, different technologies such as XR technologies (Mystakidis 2022), NFTs, hardware (for infrastructure) and 3D virtualization engines were identified (Dolata and Schwabe 2023). We found two infrastructure clusters, covering the fields of data, network and security infrastructure. We therefore can only partly confirm that the technologies mentioned in literature are developed for the Metaverse. However, it is possible that technologies for the metaverse are also developed in an application cluster if the technology is primarily used for this purpose and not as infrastructure. For instance, Oculus VR (now part of Meta), produces head mounted displays which are an XR technology. Their knowledge is associated to the "Digital Entertainment" cluster, as this is a technology not related to Metaverse infrastructure but to specific entertainment related applications. A further explanation could be that the Metaverse strongly builds upon generalistic digital infrastructures, "comprising an installed base of diverse information technology capabilities and their user, operations, and design communities" (Tilson et al. 2010, p. 748-749). Accordingly, no new knowledge generation is required for the basic technological foundation. For instance, powerful hardware is required (e.g. Dincelli and Yayla (2022) and (Dolata and Schwabe 2023)), but the development of these is not Metaverse-specific as manufacturers work on hardware for several purposes like artificial intelligence and cloud computing. As our analysis suggests, however, actors build Metaverse-specific knowledge in very selective infrastructural areas and for very different application scenarios, supported by the fact that all clusters are characterized by a variety of CPC codes.

Actor Perspective

The actor-related cluster analysis revealed seven firm clusters. We find that, although firms within the clusters are very similar in their knowledge generation practices, they differ substantially in terms of their origins and affiliations. Our findings show a vast variety of SIC codes in each firm cluster, highlighting the diversity of actors with heterogeneous knowledge and capabilities from different ecosystems. For instance, the fourth cluster comprises, on the one hand, Oculus VR (today part of Meta), a producer of head mounted displays and, on the other hand, Warner Bros. Entertainment Inc, a company from the cinema business. This finding lends support to current literature suggesting that the Metaverse can be viewed as a meta-ecosystem connecting multiple ecosystems (Schöbel and Leimeister 2023). Our results also suggest that the Metaverse evolves via inter-ecosystem interactions, that is, actors from one ecosystem can participate in another ecosystem and contribute with their capabilities to this ecosystem (Biedebach et al. 2021) and do not rely on only one ecosystem (Selander et al. 2013). IS literature further indicates digital innovations require heterogeneous knowledge (Yoo et al. 2012) making it accordingly a crucial factor. Aside from the IS literature authors noted that the building of nascent industries is driven by the knowledge creation and aggregation of actors (Moeen et al. 2020) and that different knowledge is important for the Metaverse (Dincelli and Yayla 2022). By highlighting the variety of industries actors come from we can state that the knowledge base is already very heterogeneous, and firms contribute heterogeneous knowledge and capabilities. We also believe that the vast number of firms and heterogeneous knowledge base also indicates that it may come to disagreements of actors seeking to define standards in the Metaverse, as was found for other nascent ecosystems like the mobile payment ecosystems (Du 2018). While every firm cluster contains several different SIC codes, it can be observed that computer programming and data processing-related SIC codes, which usually relate to software, are present in every firm cluster. This indicates that in each firm cluster the capabilities and knowledge of firms from these industries can be found making this industry sector a substantial part of the Metaverse knowledge base. This aspect can be found implicitly in literature: Authors usually highlight the importance of technologies related to software like NFTs and 3D virtualization engines (Dolata and Schwabe 2023; Schöbel and Leimeister 2023). Nonetheless they back our findings: A lot of software related firms will be part of the Metaverse and seem to have either application specific knowledge or general technical knowledge, as they contribute their knowledge to both or several types of found technology areas.

Combined Perspective

Prior literature highlighted that digital technology and associated actors constitute digital innovation ecosystems (Nischak et al. 2017; Wang 2021). We investigated both by performing cluster analyses for CPC codes and based a second cluster analysis on the results to identify firm clusters. Combining the results of the technology and firm clustering, we find that firms seem to pursue different knowledge generation strategies when it comes to the Metaverse. A closer look at the firm and technology clusters reveals firms that seem to "fit" the associated technology cluster very well while others do not. For instance, Leica Microsystems CMS produces microscopes and is in the firm cluster in which firms primarily have patents in the "Healthcare" cluster. OSRAM, primarily a manufacturer of lighting equipment, is located in cluster 1 with companies holding primarily patents in the technology cluster "Network and Security Infrastructure". This shows that OSRAM has created knowledge in a technology area that goes beyond their usual business, while Leica Microsystems CMS has created knowledge in the area they know. Authors noted that firms start to evaluate how the Metaverse affects them and already started their Metaverse initiatives (Peukert et al. 2022). This requires either new knowledge or the adaptation of existing knowledge. Researchers identified different generic knowledge strategies applied by firms, for instance Exploiters, with low R&D spending and the aim to increase the benefit of a single product by finding new use cases or Innovators, who learn fast and offer radically new products (Bierly and Chakrabarti 1996). We assume that firms in the Metaverse also follow such generic knowledge strategies: Our results suggest that different knowledge strategies probably are applied. It seems that some firms focus their knowledge on areas, they are familiar with, while other follow a more diversified strategy. As the above example shows, we identified firms that create knowledge in their usual business, firms that go beyond that and create knowledge in areas new for them and further those firms probably doing both, as they create knowledge in different technology areas.

Another nuance in the knowledge generation strategies applied by firms is that some actors create knowledge rather in infrastructure-related clusters, others rather in application-related clusters, which suggests that they either aim to deliver foundational technology necessary to build and operate the Metaverse or offer specific applications in the Metaverse. Further, our results revealed actors that are active in both, application and infrastructure. NVIDIA is a quite good example, as the firm is active in the semiconductor industry and therefore part of the technical infrastructure of the Metaverse, but also offers the Omniverse as an exemplary application in the Metaverse. According to IS literature digital innovations can be found on different architectural layers which are loosely coupled (Yoo et al. 2010). These layers include a device, a network, a service and a content layer. Firms can address one or several layers (Selander et al. 2013) as innovation is created across different architectural layers (Yoo et al. 2010). This can be confirmed by the results of the study. We found firm clusters either focused on infrastructure or applications and moreover a cluster with firms creating knowledge for both. Assuming that these layers can be applied to the Metaverse, we further can assume that technology clusters can be on different layers. This suggest that a firm focusing on an infrastructure cluster probably aims to create knowledge for the device or network layer, while a firm with a specific application aims to deliver knowledge for service and content layer. However, the boundaries are blurred, as application clusters also contain hardware related patents and firms, meaning the firms can be assigned to different layers depending on their patents. The example of NVIDIA further shows that companies also create knowledge for more than one layer, highlighting the aspect of innovation created across different architectural layers.

Implications

This study aims to contribute to the emerging Metaverse literature by systematically investigating the knowledge base and identifying both, actors, and applications of the Metaverse. We identified actors creating knowledge in form of patents, which are also named in literature like Adidas, NVIDIA and Microsoft (Dolata and Schwabe 2023). Further we identified Metaverse applications in our study which suggest that some of the in literature mentioned possible applications (Dincelli and Yayla 2022; Park and Kim 2022; Zahedi et al. 2022) have an existing knowledge base in the Metaverse context. While we cannot confirm the existence of all these applications, we empirically identified application and infrastructure clusters based on existing patents, while other researchers based their findings primarily on literature and examples from practice. A further contribution is the identification of a Metaverse knowledge base which is, according to literature, crucial for further development of an emerging ecosystem (Moeen et al. 2020; Moeen and Agarwal 2017; Moore 1993). With this empirical verification we can also confirm that the Metaverse is not only a marketing buzz word (Peukert et al. 2022), but has an existing knowledge base with patented knowledge. Our study further contributes to Metaverse literature by applying the ecosystem lens proposed by other authors (Schöbel and Leimeister 2023; Seidel et al. 2022). We can confirm that this perspective is valuable as we identified heterogeneous actors that originate from other ecosystems contributing to the knowledge base of the Metaverse. In extant work, the Metaverse is viewed as an ecosystem of ecosystems or meta ecosystem connecting multiple ecosystems (Schöbel and Leimeister 2023; Seidel et al. 2022), which relates to the holarchy of digital innovation ecosystems with multiple levels of ecosystem types as a possible approach for explanation (Wang 2021). We argue that the Metaverse might fits this holarchy to a certain degree and perhaps also offers an additional perspective: We found a firm cluster consisting of heterogenous firms who build knowledge in every technology cluster. While this cluster illustrates activities in different technology fields, the six remaining clusters have a clear focus on one technology cluster. We further identified six technology clusters of which four are application specific and two are infrastructure related. We argue that these six technology clusters indicate category ecosystems, while the firm cluster in which firms hold patents in each technology cluster indicates a new business/tech landscape pertaining specifically to the Metaverse. The Metaverse seems to connect existing (category) ecosystems which is in line with literature, stating that the Metaverse is a meta ecosystem connecting different existing ecosystems (Schöbel and Leimeister 2023). We further propose that with the Metaverse a virtual layer could be added to existing ecosystems: For instance, “Healthcare” is not a typical application considered when thinking about virtual reality, but it is one that can be identified in the knowledge base of the Metaverse.

Moreover, our results offer valuable insights for practitioners as we identified an existing knowledge base for Metaverse infrastructure and applications, created by heterogeneous actors. The chase for knowledge and applications started, for example, Nike, as a sporting goods manufacturer, created patents to ensure that they can sell products virtual in the Metaverse. Knowledge will likely be a key factor for successfully doing business in the virtual world of the Metaverse. Firms therefore should pro-actively check possible strategies for the Metaverse, potentially viewing our results and the identified strategies as a blueprint. They should answer the question whether the firm aims to follow a diversification strategy or a more focused strategy. To be specific, they might either focus on adapting existing knowledge for the Metaverse or on diversifying by creating new knowledge applied in new fields in the Metaverse. For some firms it might also make sense to do both, for instance by contributing to the underlying infrastructure with adapted existing knowledge and offer applications based on created new knowledge. Either way the Metaverse seems to offer various diversification opportunities that firms should scrutinize carefully. Further, firms can evaluate based on our findings which knowledge already exists and find opportunities for their own knowledge creation. We also show that it is possible to identify possible partners with existing knowledge by examining the knowledge base of the Metaverse.

Limitations and Future Research

Factors Limiting Generalizability

Our study is surely not free from limitations. We limited our patent identification efforts on the USPTO, which is obviously very biased towards the United States. However, a lot of patents are registered in the United States, even if the filing firm is from a foreign country, which means the number of patents not found is probably very low. In the USPTO database, we searched for patents explicitly referring to the Metaverse. We purposefully selected this approach to be as focused as possible on actors and technologies that explicitly address the Metaverse to be able to scrutinize the actual substance related to this phenomenon, which is surely surrounded by a lot of hype and buzz. This approach might be a limitation as there may be further patents relevant for the knowledge base of the Metaverse, which this focus failed to capture as they might utilize other terms such as “Virtual World”. The limits of a keyword-based search were also mentioned by other authors who combined it with patent classes (Battke et al. 2016), an approach that is not suitable for an explorative search in the context Metaverse, which is not related to specific patent classes. Furthermore, our approach cannot ensure that our dataset contains only patents that have already found application in the Metaverse context. Nonetheless, other approaches might be also promising, which we further outline in the Agenda for Future Research. Another limitation refers to the cluster analysis, as we decided to apply two-step clustering and other approaches might be similar appropriate. However, as a two-step cluster analysis allows to evaluate the solution for contextual fit similar to hierarchical clustering, we are convinced that the decision made sense. Further, it is possible that a different order of the two cluster analyses would yield different results. Similar to authors who used patent data, we focused on the knowledge from a structural perspective rather on details on firm level (Bierly and Chakrabarti 1996; Lou and Wu 2021; Stephan et al. 2017). It is also possible to investigate the knowledge on firm or patent level, which might generate additional insights which this study cannot deliver. Moreover, the dataset in general may contain information for additional insights, for example, patterns between objects of technology and firm clusters which we did not systematically analyzed in this study.

An Agenda for Future Research

Based on the aforementioned limitations as well as the findings of our study, we can derive an agenda for future IS research related to the knowledge generation structures in the Metaverse. This research agenda is structured along three key themes- *antecedents*, *manifestations* and *consequences* of knowledge generation structures. A first step, related to the *antecedents*, would be the identification of factors leading firms to engage in knowledge generation in particular technology areas or application scenarios as well as, ultimately, the adoption of a specific knowledge strategy. For instance, internal factors like pre-existing knowledge, spending on R&D, financial power or top management team (TMT)-structures (such as the presence of a CIO or CDO in the TMT) or external factors like competitive forces or potential cooperation partners might drive firm decisions about its Metaverse investment. A second step would be a further analysis of firms, as there might be substantial differences regarding technology fields they originate from. Some might be typical technology related firms, while others are more traditional firms which have no connection to software or hardware, but to possible fields of application instead.

Second, regarding the *manifestations* of knowledge generation structures in the Metaverse, a good future research opportunity is to refine our results. One conceivable way to improve our results is by creating a dictionary or applying machine learning approaches to find patents not using the term Metaverse and additionally ensure that all patents fit the Metaverse context. Such a study could reveal a more exhaustive list of Metaverse related patents. Another approach would be to use other sources like annual reports of companies to identify further actors. We have identified seven firm clusters which are groups of actors, but an investigation of the specific roles these actors take was not part of our study. Therefore, we suggest to investigate which roles are relevant in the Metaverse and how they can be characterized (Biedebach and Hanelt 2020). Further, there might be differences related to the approach of firms to profit from an emerging Metaverse ecosystem. For instance, there might be firms actively engaged in the Metaverse and those who create complementary knowledge. As the Metaverse is an emerging ecosystem knowledge is created rapidly. Therefore, we suggest to repeat the investigation of the knowledge base in future. It is likely that every month new knowledge is generated and patented. As soon as the knowledge is increased and more connections between patents arise based on citations, it is appropriate to apply a patent citation network following great role models in past research (Ardito et al. 2018; Stephan et al. 2017). Additionally, Han et al. (2022) identified the dimensions roles, structures and processes for multiple stages of ecosystem development. Applying their model to the Metaverse might increase the understanding by giving more insights to characteristics and possible challenges in both, the current and future stages of the ecosystem.

Third and finally, several opportunities for future work arise that relate to the *consequences* of knowledge generation in the Metaverse. On the one hand, it seems promising to evaluate the performance of the identified strategies in the Metaverse context. For instance, are there strategies better suite a specific industry than others? This would be a valuable contribution for practitioners as well as it is a starting point allowing to adapt the firm's strategy for the Metaverse. This aspect includes the question of whether companies that actively work on the Metaverse and other companies that do not but create related knowledge have different approaches to knowledge generation and knowledge strategies. A further important factor in the context of these knowledge strategies is the question whether created knowledge is valuable or not. This could be identified by using a proxy like citations of a certain patent. When firms decide to join the Metaverse, they must probably decide whether they want to be part of a firm cluster related to an application cluster, an infrastructure cluster or to several technology clusters. As this decision will influence the strategy and the knowledge generation, it would be interesting to evaluate what performance implications engaging in different clusters has. With increasing interest in the Metaverse both practitioners and researchers will probably search for opportunities of new applications, but also for existing ones that could be migrated to the Metaverse. We outlined the heterogenous actors in the Metaverse ecosystem which originated in different ecosystems. Based on this finding, a highly interesting future research opportunity is to investigate interactions of these actors on different levels of ecosystems (Wang 2021). It would be interesting to see if there are interactions and if yes what kind of interactions.

References

Adner, R. 2017. "Ecosystem as Structure: An Actionable Construct for Strategy," *Journal of Management* (43:1), pp. 39–58.

Adner, R., and Kapoor, R. 2010. "Value Creation in Innovation Ecosystems: How the Structure of Technological Interdependence Affects Firm Performance in New Technology Generations," *Strategic Management Journal* (31:3), pp. 306–333.

Ardito, L., D'Adda, D., and Messeni Petruzzelli, A. 2018. "Mapping Innovation Dynamics in the Internet of Things Domain: Evidence from Patent Analysis," *Technological Forecasting and Social Change* (136), pp. 317–330.

Baldwin, C. 2012. "Organization Design for Business Ecosystems," *Journal of Organization Design* (1:1), pp. 20–23.

Battke, B., Schmidt, T. S., Stollenwerk, S., and Hoffmann, V. H. 2016. "Internal or External Spillovers—Which Kind of Knowledge Is More Likely to Flow within or across Technologies," *Research Policy* (45:1), pp. 27–41.

Biedebach, M., and Hanelt, A. 2020. "Towards a Typology of Ecosystem Roles in the Era of Digital Innovation – An Inductive Empirical Analysis," in *Proceedings of the 41st International Conference on Information Systems (ICIS 2020)* (Vol. 7), India.

Biedebach, M., Hanelt, A., and Firk, S. 2021. "Inter-Ecosystem Migration of IT-Actors in the Era of Digital Innovation: An Empirical Investigation," in *Proceedings of the 42nd International Conference on Information Systems (ICIS 2021)*, Austin.

Bierly, P., and Chakrabarti, A. 1996. "Generic Knowledge Strategies in the U.S. Pharmaceutical Industry," *Strategic Management Journal* (17:S2), pp. 123–135.

Bobrowsky, M. 2023. "The Metaverse Is Quickly Turning Into the Meh-Taverse," *Wall Street Journal*, , March 29. (https://www.wsj.com/articles/the-metaverse-is-quickly-turning-into-the-meh-taverse-1a8dc3d0?mod=e2li, accessed May 4, 2023).

Cavanaugh, J. E., and Neath, A. A. 2019. "The Akaike Information Criterion: Background, Derivation, Properties, Application, Interpretation, and Refinements," *WIREs Computational Statistics* (11:3), pp. 1–11.

Ceccagnoli, Forman, Huang, and Wu. 2012. "Cocreation of Value in a Platform Ecosystem! The Case of Enterprise Software," *MIS Quarterly* (36:1), p. 263.

Chiang, C.-H., Liaw, Y.-C., and Wu, R.-T. 2017. *Patent Analysis for the Formulation of Technology Policy: Evidence from 4G LTE Industry*, presented at the Hawaii International Conference on System Sciences.

Clarysse, B., Wright, M., Bruneel, J., and Mahajan, A. 2014. "Creating Value in Ecosystems: Crossing the Chasm between Knowledge and Business Ecosystems," *Research Policy* (43:7), pp. 1164–1176.

Costantini, V., Crespi, F., and Curci, Y. 2015. "A Keyword Selection Method for Mapping Technological Knowledge in Specific Sectors through Patent Data: The Case of Biofuels Sector," *Economics of Innovation and New Technology* (24:4), pp. 282–308.

Dattée, B., Alexy, O., and Autio, E. 2018. "Maneuvering in Poor Visibility: How Firms Play the Ecosystem Game When Uncertainty Is High," *Academy of Management Journal* (61:2), pp. 466–498.

Diaz, J. 2023. "BMW's New Factory Doesn't Exist in Real Life, but It Will Still Change the Car Industry," *Fastcompany.Com*, , March 21. (https://www.fastcompany.com/90867625/bmws-new-factory-doesnt-exist-in-real-life-but-it-will-still-change-the-car-industry, accessed May 4, 2023).

Dincelli, E., and Yayla, A. 2022. "Immersive Virtual Reality in the Age of the Metaverse: A Hybrid-Narrative Review Based on the Technology Affordance Perspective," *The Journal of Strategic Information Systems* (31:2), pp. 1–22.

Dionisio, J. D. N., Iii, W. G. B., and Gilbert, R. 2013. "3D Virtual Worlds and the Metaverse: Current Status and Future Possibilities," *ACM Computing Surveys* (45:3), pp. 1–38.

Dolata, M., and Schwabe, G. 2023. "What Is the Metaverse and Who Seeks to Define It? Mapping the Site of Social Construction," *Journal of Information Technology* (0:0).

Du, K. 2018. "Complacency, Capabilities, and Institutional Pressure: Understanding Financial Institutions' Participation in the Nascent Mobile Payments Ecosystem," *Electronic Markets* (28:3), pp. 307–319.

Forman, C., and van Zeebroeck, N. 2011. "From Wires to Partners: How the Internet Has Fostered R&D Collaborations within Firms," *Management Science* (58:8), pp. 1549–1568.

Haki, K., Blaschke, M., Aier, S., Winter, R., and Tilson, D. 2022. "Dynamic Capabilities for Transitioning from Product Platform Ecosystem to Innovation Platform Ecosystem," *European Journal of Information Systems*, pp. 1–19.

Han, J., Zhou, H., Lowik, S., and De Weerd-Nederhof, P. 2022. "Enhancing the Understanding of Ecosystems under Innovation Management Context: Aggregating Conceptual Boundaries of Ecosystems," *Industrial Marketing Management* (106), pp. 112–138.

Hanelt, A., Firk, S., Hildebrandt, B., and Kolbe, L. M. 2021. "Digital M&A, Digital Innovation, and Firm Performance: An Empirical Investigation," *European Journal of Information Systems* (30:1), pp. 3–26.

Hodapp, D., Hawlitschek, F., and Kramer, D. 2019. "Value Co-Creation in Nascent Platform Ecosystems: A Delphi Study in the Context of the Internet of Things," *Proceedings of the 40th International Conference on Information Systems (ICIS), Munich*.

Huang, P., Ceccagnoli, M., Forman, C., and Wu, D. J. 2009. "When Do ISVs Join a Platform Ecosystem? Evidence from the Enterprise Software Industry," in *Proceedings of the 13th International Conference on Information Systems (ICIS 2009)* (Vol. 161), Phoenix.

Huang, P., Tafti, A., and Mithas, S. 2018. "Platform Sponsor Investments and User Contributions in Knowledge Communities: The Role of Knowledge Seeding," *MIS Quarterly* (42:1), pp. 213–240.

IBM. 2021. "Model Summary View," *IBM.Com*, , March 22. (https://www.ibm.com/docs/en/spss-statistics/25.0.0?topic=viewer-model-summary-view, accessed May 4, 2023).

IBM. 2022. "TwoStep Cluster Analysis," *IBM.Com*, , October 5. (https://www.ibm.com/docs/en/spss-statistics/saas?topic=features-twostep-cluster-analysis, accessed May 4, 2023).

Jacobides, M. G., Cennamo, C., and Gawer, A. 2018. "Towards a Theory of Ecosystems," *Strategic Management Journal* (39:8), pp. 2255–2276.

Kaufman, L., and Rousseeuw, P. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*, Hoboken, New Jersey, USA: John Wiley & Sons, Inc.

Kent, P., Jensen, R. K., and Kongsted, A. 2014. "A Comparison of Three Clustering Methods for Finding Subgroups in MRI, SMS or Clinical Data: SPSS TwoStep Cluster Analysis, Latent Gold and SNOB," *BMC Medical Research Methodology* (14:1), pp. 1–14.

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

Moeen, M., and Agarwal, R. 2017. "Incubation of an Industry: Heterogeneous Knowledge Bases and Modes of Value Capture," *Strategic Management Journal* (38:3), pp. 566–587.

Moeen, M., Agarwal, R., and Shah, S. K. 2020. "Building Industries by Building Knowledge: Uncertainty Reduction over Industry Milestones," *Strategy Science* (5:3), pp. 218–244.

Moore, J. F. 1993. "Predators and Prey: A New Ecology of Competition," *Harvard Business Review* (May-June 1993), pp. 75–86.

Mystakidis, S. 2022. "Metaverse," *Encyclopedia* (2:1), pp. 486–497.

Nambisan, S. 2017. "Digital Entrepreneurship: Toward a Digital Technology Perspective of Entrepreneurship," *Entrepreneurship Theory and Practice* (41:6), SAGE Publications Inc, pp. 1029–1055.

Nischak, F., Hanelt, A., and Kolbe, L. M. 2017. "Unraveling the Interaction of Information Systems and Ecosystems - A Comprehensive Classification of Literature," in *Proceedings of the 38th International Conference on Information Systems (ICIS 2017)* (Vol. 20), Seoul.

Norušis, M. J. 2011. *IBM SPSS Statistics 19 Advanced Statistical Procedures Companion*, Prentice Hall.
Okazaki, S. 2007. "Lessons Learned from I-Mode: What Makes Consumers Click Wireless Banner Ads?," *Computers in Human Behavior* (23:3), pp. 1692–1719.
Park, S.-M., and Kim, Y.-G. 2022. "A Metaverse: Taxonomy, Components, Applications, and Open Challenges," *IEEE Access* (10), pp. 4209–4251.
Peukert, C., Weinhardt, C., Hinz, O., and van der Aalst, W. M. P. 2022. "Metaverse: How to Approach Its Challenges from a BISE Perspective," *Business & Information Systems Engineering* (64:4), pp. 401–406.
Rundle-Thiele, S., Kubacki, K., Tkaczynski, A., and Parkinson, J. 2015. "Using Two-Step Cluster Analysis to Identify Homogeneous Physical Activity Groups," *Marketing Intelligence & Planning* (33:4), pp. 522–537.
Schöbel, S. M., and Leimeister, J. M. 2023. "Metaverse Platform Ecosystems," *Electronic Markets* (33:12), pp. 1–12.
Seidel, S., Berente, N., Nickerson, J., and Yepes, G. 2022. "Designing the Metaverse," in *Proceedings of the 55th Hawaii International Conference on System Sciences (2022)*, Hawaii, pp. 6699–6708.
Selander, L., Henfridsson, O., and Svahn, F. 2010. "Transforming Ecosystem Relationships in Digital Innovation," in *Proceedings of the 31st International Conference on Information Systems (ICIS 2010)* (Vol. 138), St. Louis.
Selander, L., Henfridsson, O., and Svahn, F. 2013. "Capability Search and Redeem across Digital Ecosystems," *Journal of Information Technology* (28:3), SAGE Publications Ltd, pp. 183–197.
Stephan, A., Schmidt, T. S., Bening, C. R., and Hoffmann, V. H. 2017. "The Sectoral Configuration of Technological Innovation Systems: Patterns of Knowledge Development and Diffusion in the Lithium-Ion Battery Technology in Japan," *Research Policy* (46:4), pp. 709–723.
Stephenson, N. 1992. *Snow Crash*, Bantam Spectra Book, Bantam Books.
Tilson, D., Lyytinen, K., and Sørensen, C. 2010. "Research Commentary —Digital Infrastructures: The Missing IS Research Agenda," *Information Systems Research* (21:4), pp. 748–759.
Tkaczynski, A. (ed.). 2017. "Segmentation Using Two-Step Cluster Analysis," *Segmentation in Social Marketing*, pp. 109–125.
USPTO. 2023a. "Electronic Priority Document Exchange (PDX) Program," *USPTO.Gov*. (https://www.uspto.gov/patents/basics/international-protection/electronic-priority-document-exchange-pdx#:~:text=The%20United%20States%20Patent%20and%20Trademark%20Office%20%28USPTO%29,foreign%20priority%20applications%20as-filed%20from%20the%20participating%20offices., accessed May 4, 2023).
USPTO. 2023b. "Patent Cooperation Treaty," *USPTO.Gov*. (https://www.uspto.gov/patents/basics/international-protection/patent-cooperation-treaty, accessed May 4, 2023).
Wacker, A. K., Jurisch, M. C., Jin, J., Wolf, P., and Krcmar, H. 2014. *Identifying Main User Groups for Green IS – An Empirical Study of Electric Vehicles in China*, presented at the Thirty Fifth International Conference on Information Systems, Auckland.
Wang, P. 2021. "Connecting the Parts with the Whole: Toward an Information Ecology Theory of Digital Innovation Ecosystems," *MIS Quarterly* (45:1), pp. 397–422.
Whelan, R., and Flint, J. 2023. "Disney Eliminates Its Metaverse Division as Part of Company's Layoffs Plan," *Wall Street Journal*, , March 28. (https://www.wsj.com/articles/disney-eliminates-its-metaverse-division-as-part-of-companys-layoffs-plan-94b03650?mod=e2li, accessed May 4, 2023).
Xi, N., Chen, J., Gama, F., Riar, M., and Hamari, J. 2022. "The Challenges of Entering the Metaverse: An Experiment on the Effect of Extended Reality on Workload," *Information Systems Frontiers*.
Yoo, Y., Boland, R. J., Lyytinen, K., and Majchrzak, A. 2012. "Organizing for Innovation in the Digitized World," *Organization Science* (23:5), pp. 1398–1408.
Yoo, Y., Henfridsson, O., and Lyytinen, K. 2010. "Research Commentary —The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research," *Information Systems Research* (21:4), pp. 724–735.
Zahedi, F. M., Zhao, H., Sanvanson, P., Walia, N., Jain, H., and Shaker, R. 2022. "My Real Avatar Has a Doctor Appointment in the Wepital: A System for Persistent, Efficient, and Ubiquitous Medical Care," *Information & Management* (59:8), pp. 1–25.