Big data augmentated business trend identification: the case of mobile commerce

Ozcan Saritas[1] · Pavel Bakhtin[1] · Ilya Kuzminov[1] · Elena Khabirova[1]

Abstract

Identifying and monitoring business and technological trends are crucial for innovation and competitiveness of businesses. Exponential growth of data across the world is invaluable for identifying emerging and evolving trends. On the other hand, the vast amount of data leads to information overload and can no longer be adequately processed without the use of automated methods of extraction, processing, and generation of knowledge. There is a growing need for information systems that would monitor and analyse data from heterogeneous and unstructured sources in order to enable timely and evidence-based decision-making. Recent advancements in computing and big data provide enormous opportunities for gathering evidence on future developments and emerging opportunities. The present study demonstrates the use of text-mining and semantic analysis of large amount of documents for investigating in business trends in mobile commerce (m-commerce). Particularly with the on-going COVID-19 pandemic and resultant social isolation, m-commerce has become a large technology and business domain with ever growing market potentials. Thus, our study begins with a review of global challenges, opportunities and trends in the development of m-commerce in the world. Next, the study identifies critical technologies and instruments for the full utilization of the potentials in the sector by using the intelligent big data analytics system based on in-depth natural language processing utilizing text-mining, machine learning, science bibliometry and technology analysis. The results generated by the system can be used to produce a comprehensive and objective web of interconnected technologies, trends, drivers and barriers to give an overview of the whole landscape of m-commerce in one business intelligence (BI) data mart diagram.

✉ Pavel Bakhtin
   pbakhtin@hse.ru

Extended author information available on the last page of the article

Introduction

Mobile commerce (m-commerce) has been evolving for several decades with an increasing intensity. The first example of m-commerce was the use of mobile SMS-messages to pay for various services, such as vending machines (Scharl et al. 2005), banking (Jamil and Mousumi 2008), downloadable ringtones (Fong and Lai 2005) as well as parking, theater and cinema tickets among the others. Later, with the rapid development of online retail services, buying goods on the Internet has become a major trend with a possibility of reduced prices due to direct connection of many producers and exporters with consumers (Groß 2015). Innovation in mobile technologies and introduction of various mobile applications allowed performing online operations with the help of any wireless device, such as a smartphone or tablet PC. Since then, mobile devices have turned from being only the means of communication into personal assistants, which help accessing new information, spending leisure time, and finally, instantly purchasing digital content on demand as well as various other products and services (Nair and Bhattacharyya 2019). With the recent introduction of online mobile banking, and integration of various electronic wallet services and mobile applications (Harris and Wonglimpiyarat 2019) the popularity of m-commerce and its convenience has increased substantially (Hossain 2019). Today m-commerce is a large technology and a business area with ever growing market potentials. This trend has been triggered enormously with the measures of social isolation due to the global COVID-19 pandemic (Chiou and Tucker 2020; Okyere et al. 2020; Hasanat et al. 2020).

The present paper focuses on the emerging and evolving trends of m-commerce development. In this study, m-commerce is considered in accordance with Hew's (2017) conceptualization, which includes the following m-commerce application groups:

- Mobile (mobile value-added) services
- Mobile payment
- Mobile banking
- Mobile advertising
- Mobile applications
- Mobile internet
- Mobile shopping

Within these groups, a particular emphasis will be given to the trends in B2C and C2C segments of the market, which will cover areas such as mobile payments, mobile sales, shopping as well as new technology and topics like data security and privacy. Trends in these domains will be analyzed by using an innovative approach drawing on advanced text-mining techniques and intelligent Big Data analytics within a wide variety of data from different media sources.

Rapidly growing amount of data generated worldwide nowadays provides a good evidence for identifying emerging and future trends (Ena et al. 2016). IDC (2018) has projected that the world's total amount of data will grow from 33 to 175 zettabytes (or a billion terabytes) between 2018 and 2025, and 30% of that data will be real-time by 2025. In the future, it is expected that the textual segments of this market, including eMagazine and ePaper, will considerably expand. Their revenue is expected to show an annual growth rate (CAGR) between 2018 and 2022 of 10.8% and 7.8%, respectively (IDC 2017). Analysis of these growing volumes of information with the use of traditional methods is a serious challenge for evidence-based and timely decision-making. Application of the Big Data analysis
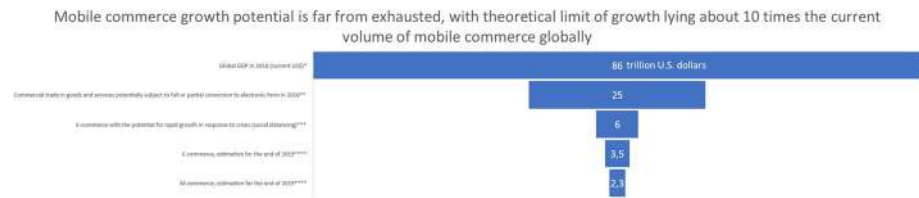
Fig. 1 Market potential of global m-commerce. *Sources*: *The World Bank (2020), **World Trade Organization (2019a), ***World Trade Organization (2019b), ****Loesche (2018)

can dramatically increase the speed and accuracy as well as credibility of trends, forecasts, strategies, and eventual policy recommendations. The present study demonstrates the use of text-mining and semantic analysis of large amount of documents for investigating in business trends in mobile commerce. Using the 'intelligent FOResight Analytics' (iFORA) system based on in-depth natural language processing utilizing text-mining, machine learning, science bibliometry and technology analysis current study identifies critical technologies and instruments for the full utilization of the potentials in the sector. As a result we produced a comprehensive and objective web of interconnected technologies, trends, drivers and barriers to give an overview of the whole landscape of m-commerce in one business intelligence (BI) data mart diagram.

Building upon this background, the "Literature review" section of the paper begins with a review of the m-commerce as a fast growing business domain. Next, the usage and analysis of various data sources for business analysis and marketing are discussed with a particular focus on the m-commerce domain. The paper then elaborates on the methodology used for identifying sectoral trends in the "Methodology" section. The study undertaken and results generated are presented and discussed in the "Findings" section. Finally, the "Conclusions and discussion" section draws overall conclusions from the work done and discusses implications and future opportunities and threats for society, technology and policy.

Literature review

M-commerce is a subset of electronic commerce (e-commerce) and represents any electronic transaction with monetary value done through a mobile network (Ngai and Gunasekaran 2007). Total mobile e-commerce yearly sales were projected to amount to 2.32 trillion U.S. dollars in 2019 and to reach 3.56 trillion U.S. dollars in 2021 (Loesche 2018). However, these estimates should be reviewed upward as COVID-19 pandemic is poised to accelerate the growth of remote digital communications services (Petrock 2020). Trend of slowing relative growth rates of m-commerce (from 40% per year in 2016–2017 to forecasted previously 25% in 2019–2020) will also likely to reverse in recent years driven by more demand for biosecurity and remote communication and will again see annual growth of up to 40% per year for nearest future. This will lead to further increase of the already significant share of m-commerce in total e-commerce turnover (Fig. 1). This share has grown from 52.4% to 67.2% in the period 2016–2019 with the forecast of reaching up to 73% in 2021 (Loesche 2018). However, level of 75–80% by 2022 can now be

seen as probable based on the trend of developing situation around COVID-19 and global wave of nationwide lockdowns which might continue even deep into 2021.[1]

The main drivers of the m-commerce development are information technology (IT) and telecommunication technologies (TCT), as well as constantly growing popularity of online services, smartphones and mobile applications. According to Hootsuite and We Are Social (2019) report, today, there are more than four billion people, in other words, over half of the world's population, use the Internet; and two-thirds of the world's 7.6 billion inhabitants now have a mobile phone with a bulk of them being a smartphone. Mobile devices account for 48% of total Internet time and on average, the world's Internet users spend 6 h and 42 min online each day (ibid). Global m-commerce has outstanding potentials due to the widespread implementation of mobile and wearable technologies, as well as growing Internet coverage enabled by mobile Internet and free Wi-Fi connection in public places.

The main motivation for the present study to a greater extend comes from a particular research bias in the mobile domain topics discovered by Pousttchi et al. (2015) and later fortified by Tang (2019). Examining m-commerce trends in recent research papers, Pousttchi et al. (2015) found a poor coverage of many unique aspects of mobile systems (for example, real-time interactions and all their consequences for the sector development). Moreover, they concluded that much scientific work in this sphere (according to papers submitted to high-level international conferences) is done based on outdated methodologies developed formerly for information systems in general, so this research rarely can be found in top journals that demand original theoretical and methodological contributions. Even four years later, Tang (2019) pointed out that the streams of research on m-commerce are still highly inconsistent and fragmented.

Because of these blind spots, it becomes valuable to develop approaches that allow conducting an analysis with up-to-date research in highly dynamic fields to address complex business needs with strong scientific input. To overcome described constraints of traditional research, and meet the aforementioned scientific and business requirements, the text mining, and more lately tech-mining, approaches have been introduced into social, economic and science and technology (S&T) research (Porter and Cunningham 2004). Generally, these studies involve a combination of bibliometrics, patent analysis and different Natural Language Processing (NLP) tools to collect, process and represent competitive technological intelligence in a visual form.

Due to the fact that many scientific articles focus on research, they may significantly lack information about technology and market trends at the industry-level. Other sources of information can be used to gain technology and market intelligence. For instance, professional sectoral media sources can be useful for up-to-date business analytics and marketing research. Broader media and news sources are often used when an analysis has a goal to identify on-the-spot patterns and deviations like criminal networks activity (Pramanik et al. 2017) or stock price fluctuations due to newsworthy events (Wang et al. 2019, Nassirtoussi et al. 2015). Various monitoring research can benefit from constructing an automated tracking of text mining indicators based on digital media data, such as sustainability indicators developed by Juventia et al. (2020) and Rivera et al. (2014), or economic policy uncertainty index with high predictive power proposed by Yono et al. (2019) and Tobback et al. (2018).

---

[1] For more estimates on m-commerce market see Petrov C. (2019). 50 M-Commerce Statistics for 2020. https://techjury.net/stats-about/mcommerce/. (Last accessed date 7 April 2020).

This type of data source possesses unique features such as top-of-the-agenda and widespread coverage of all the factual aspects of almost any significant issue, including multidisciplinary relations and various types of trends, for example in social, technological, economic, ecological, political and value (STEEPV) systems (Miles et al. 2016). More complex text mining approaches include among others the topic evolution maps based on the cross-media corpus analysis (see the recent review of this research field in Zhou et al. 2017) and comparison of media and scientific discourses about particular topic (Kayser 2017).

Therefore, preparing representative sampling of sectoral media information sources for Big Data and text mining processing can be beneficial, given that media data has also some limitations. For example, media and websites rarely cover information about proprietary technologies and particular names of techniques and instruments, while concerning more abstract objects and entities (like in studies by Chiarello et al. 2019, Gök et al. 2015).

Until recently, simple isolated bibliometrics and patent analysis methods based on official classifications and citation indexes dominated the technology and industry research and horizon scanning studies (Bakhtin et al. 2017). Frequently, researchers tend to scope their field of study on clearly defined small samples of documents. The limitation of such approaches lies in the fact that it is impossible to evaluate whether enough text corpus is analyzed to make a valid inference due to lack of confidence about data boundaries for fields. Traditional orientation of many researches towards structured metadata highly pertains to methodological, computational and technical difficulties of integrated text mining. Using existing classifications not only prevents researchers from catching the young areas of innovation and science, yet not included in official taxonomies, but prevents them also from optimally setting topic boundaries, as opposed to modern methods of topic clustering/topic modelling approaches, which have become a de-facto standard in text mining applications since the mid-2000s (Chuang et al. 2013; Zhang et al. 2014; Moro et al. 2019). The approach based on bare keywords without semantic multipliers (such as word2vec based associative/synonymous series of terms) is even weaker than the use of official taxonomies, as the results are ultra-sensitive to meaningless fluctuations in word usage patterns, homonyms, and lexical hypes (see application of these approaches in Goldberg and Levy 2014; Rong 2014; Vidra 2015; Kim et al. 2020). To find, download and process the most relevant documents that potentially contain information about emerging S&T areas, researchers apply iterative processes from general search categories and terms to more concrete words and phrases (Ju et al. 2020; Huang et al. 2015).

Taking into account these numerous drawbacks of such linear, non-machine learning (ML) approaches along with an absence of complex Big Data sources, the present research is mainly concerned with the question of: how Big Data augmented horizon scanning can contribute to business trends analysis. Answer to this question starts with a detailed description of methodology applied. The proposed methodology is then demonstrated with the case of the m-commerce domain. Particularly with the on-going COVID-19 pandemic and resultant social isolation, m-commerce has become a large technology and business domain with ever growing market potentials. So, there is a growing need for monitoring and analysis of data about the sector from heterogeneous and unstructured sources in order to enable timely and evidence-based decision-making.
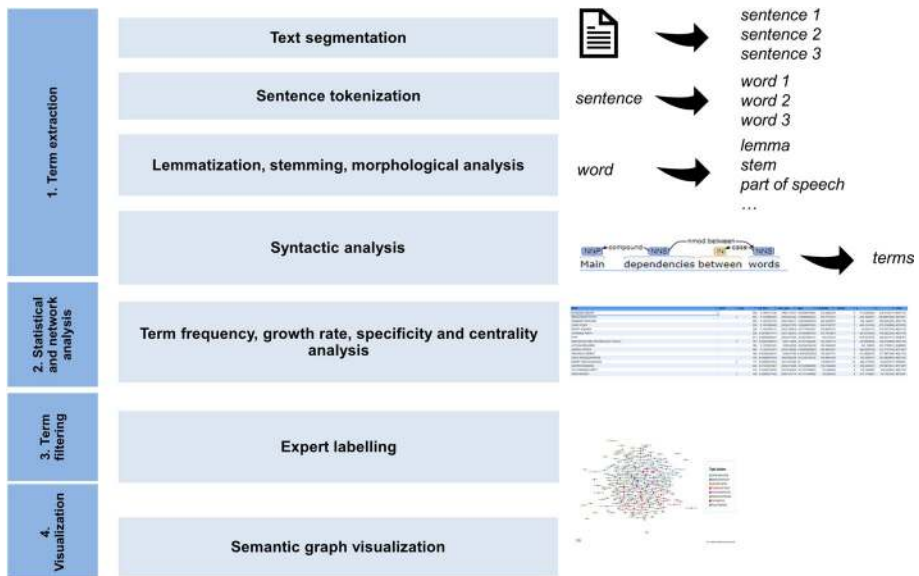
Fig. 2 Stages and methods of analysis of the collected text array and related metadata. *Source*: HSE ISSEK

Methodology

The process of the identification of global trends in m-commerce was carried out using the 'intelligent FOResight Analytics' (iFORA) system, which was developed at Institute for Statistical Studies and Economics of Knowledge (ISSEK), National Research University Higher School of Economics (HSE), Moscow.[2] This intelligent big data analytics system is based on an in-depth natural language processing with new algorithmic developments as well as the application of the best practices of text-mining, machine learning, science bibliometry and technology analysis. iFORA draws on an integrated database of diverse information and data sources. The iFORA database (DB) hosts more than 350 million documents, including 40 million international news reports, 10 million Russian news reports, 120 million scientific publications in all research areas, 100 million patents, as well as grant applications and grants, analytical and strategic documents and other sources of information.

Global trends in m-commerce that influence the development of the area in the medium and long term were identified as a result of a syntactic-semantic analysis. As we showed at the beginning of the paper, media is a valuable information source for business analysis and marketing, especially in the m-commerce domain. For this reason the sources of data used for this analysis included reports in the field of STI development and venture capital. A search in the top ranking news portals with science and technology focus in Alexa.com and SimilarWeb. com yielded about 200 world leading media sources for the analysis including MIT Technology Review, businessinsider.com, venturebeat.com, techcrunch.com, among the others. The time period for the analysis was 2013 to 2019, during when a significant development of the m-commerce sector was observed. Data collection from a variety of open sources was carried

---

[2] URL: https://issek.hse.ru/en/ (date last accessed 12/06/2020).

out by parallel (multi-threaded and distributed) launch of software data extraction processes. Relevant texts for the analysis were extracted with the following search query: ***"m-commerce" OR "mobile-commerce" OR ("mobile" AND ("commerce" OR "e-commerce")).***

The procedure of processing and analysis of text data collected consisted of four main stages: (1) *term extraction,* (2) *statistical and network analysis,* (3) *term filtering,* and (4) *visualization.* The process is illustrated in Fig. 2.

At the first stage of *term extraction*, large sets of unstructured textual data were transformed into structured tables, arrays and vector representations. For this purpose, each document was split into separate sentences, words, and phrases with different linguistic characteristics. Syntactic analysis included an in-depth assessment of the text in order to identify its semantic meaning on the basis of syntactic structure of the links between words in the sentence. For all words in the sentence, the chains of word dependencies (where the first word is the governor, the second is dependent) were identified as well as the type of communication depending on the context of their interaction. The following dependencies were considered as candidates for terms: ⟨*adjective, adjective modifier, noun*⟩ (e.g. social commerce), ⟨*noun, compound, noun*⟩ (e.g. m-commerce), ⟨*noun, noun modifier, noun*⟩ (e.g. internet of things). The procedure was done iteratively combining unigrams into bigrams, bigrams into trigrams and so on.

Once terms were identified and mapped to sentences in respective documents, the statistical analysis was applied. Five main indicators were calculated:

- Relative term frequency,
- Term growth rate,
- Term specificity,
- Term centrality,
- Term score.

The annual relative term frequency $tf_i(t)$ was estimated using following formula (1):

$$tf_i(t) = \frac{\left|\left\{s_i \in S_i | t \in s_i\right\}\right|}{|S_i|},$$ (1)

$\left|\left\{s_i \in S_i | g \in s_i\right\}\right|$—the amount of times the term t occurs in the sentences in all documents (not only m-commerce) of a ith publication year,$|S_i|$—the amount of all occurrences in all sentences in all documents of the ith publication year.

The total relative term frequency $tf(t)$ was assessed using the following formula (2):

$$tf(t) = \frac{|\{s \in S | t \in s\}|}{|S|},$$ (2)

which is almost the same equation as (1) but for all years of study.

The following indicator helps to assess the global significance of the topic in the whole database.

Term growth rate $gr(t)$ was calculated using the following formula (3):

$$gr(t) = \frac{1}{T}\left(\frac{\sum_{i \geq T/2} tf_i(t)}{\sum_{i \leq \frac{T}{2}} tf_i(t)} - 1\right),$$ (3)

T—the length of time period in years, $i = 1, 2, \ldots, T$, $tf_i(t)$—annual relative term frequency in ith publication year.

Term growth rate allows spotting trends in the data—the higher the overall dynamics, the more potential this topic may gain in time.

In order to help distinguishing between relevant and irrelevant terms, we calculated specificity indicator that would compare domain (m-commerce related) relative term frequency against total relative term frequency in all documents using following formula (4):

$$specificty\ (t) = log10\frac{tf(t)_{domain}}{tf(t)}, \tag{4}$$

$tf(t)_{domain}$—relative term frequency in documents D_{domain}, where query *("m-commerce" OR "mobile-commerce" OR ("mobile" AND ("commerce" OR "e-commerce")))* is true for every document $d \in D_{domain}$.

The values of $specificty(t) > 0$ accounts for terms, which proportionally occur in m-commerce documents more often than in the whole iFORA database, making them relatively more relevant.

Next step was centrality calculation. In order to construct the graph G, we denoted V as the set of the terms (graph nodes) and E, as the set of edges between the nodes, so that $G = \langle V, E, w \rangle$, where $w : E \to \mathbb{R}_+^*$—is the edge weight function determined by co-occurrence relationship.

We defined co-occurrence relationship as an existence of connection between two terms in the window context size of k sentences within each document to count links between terms. To calculate edge weight, we used following formula (5):

$$w(x, y) = \sum_{d \in D, |m-n| \leq k, m \in s(x,d), n \in s(y,d), x \neq y} 1 \tag{5}$$

D—set of all documents, d—a document, $d \in D$, x—first term, $x \in V$, y—second term, $y \in V$, $s(x, d)$, $s(y, d)$—functions that return lists of sentence numbers where terms x and y respectively occur in a document d, m, n—sentence numbers, $m \in s(x, d), n \in s(y, d), k$—a window context size.

For our study we picked $k = 9$.

Once co-occurrence was estimated and the graph was formed, the term *centrality* was calculated for each term x using the following formula (6) (inspired by Opsahl et al. 2010):

$$centrality(t) = \sqrt{\sum_{y \in V} w(t, y) * \sum_{y \in V, w(t,y)>0} 1}, \tag{6}$$

x—is a term for which the calculation is being estimated, $w(t, y)$—co-occurrence between term t and y. The higher centrality accounted for more connected terms showing their impact in the network.

By aggregation of all four indicators and removing terms with $specificty(t) \leq 0$ and $gr(t) \leq 0$, the final *score* indicator was calculated as a measure of the general informative value and importance of each term in documents processed. The following formula (7) was used to normalize values between 1 and 5:

$$score(t) = \left(\frac{tf(t) - \min(tf)}{\max(tf) - \min(tf)} * 4 + 1 \right)^{\frac{1}{4}} * \left(\frac{gr(t) - \min(gr)}{\max(gr) - \min(gr)} * 4 + 1 \right)^{\frac{1}{4}}$$
$$* \left(\frac{specificity(t) - \min(specificity)}{\max(specificity) - \min(specificity)} * 4 + 1 \right)^{\frac{1}{4}} * \left(\frac{centrality(t) - \min(centrality)}{\max(centrality) - \min(centrality)} * 4 + 1 \right)^{\frac{1}{4}} \tag{7}$$

The results of statistical analysis helped to filter out insignificant topics—terms with exceptionally low *score*. However, due to the fact that we wanted to spot relevant m-commerce topics, an additional filtering was done to get rid of stopwords, country and people names and other irrelevant topics that still happened to get relatively high *score*.

In future research we plan to use results of such expert filtering to train supervised machine learning models to spot relevant topics in texts in automated way. To achieve that, it is planned to combine word embedding techniques, such as *word2vec/sense2vec* (Trask et al. 2015) and *Named Entity Recognition (NER) model* as part of *convolutional neural network*. While not enough labeled data is available for research, the expert judgement is left as the main filtering method.

Once the analysis was complete, results were visualized in the format that would be intuitive and insightful for potential decision makers, researchers, and experts. To accomplish that we used the *semantic graph* instrument.

Semantic graph is a graph of term co-occurrences. We applied Louvain community detection algorithm created by Blondel et al. (2008) to cluster terms into topics based on co-occurrences and represent each node size by *score* indicator. Color of each node was specified by its affiliation to the cluster.

Based on visual outputs, which are the results of aggregation of large analytical tables containing lists of identified terms ranked based on combination of text-mining metrics, the main global trends and prospective directions for development associated to m-commerce were discovered.

Findings

Earlier Pousttchi et al. (2015) found a way of structuring the empirically derived m-commerce landscape through the framework of "nomadic information environments", which was defined by its authors as "heterogeneous assemblage of interconnected technological, and social, and organizational elements that enable the physical and social mobility of computing and communication services between organizational actors both within and across organizational borders" (Lyytinen and Yoo 2002, p. 377). The approach developed in the present study for data analysis and visualization allows representing clearly the complex interconnections within m-commerce sector. In order to illustrate these complex and systemic relationships, a semantic graph was constructed (Fig. 3). The graph shows the thematic structure of the m-commerce domain with its narrow themes, directions of development and their interconnections with certain clusters.

The nodes of the graph (points on the semantic graph in Fig. 3) are the terms that are revealed by the results of the machine analysis of the array of full texts; and the edges are the links between the terms. Each term may reflect a separate direction of development, trend, technology, market, product or other meaningful concept. Machine-generated *clusters* of terms are marked with different colors and numbered in Fig. 3. They are formed in such a way that between the terms inside each cluster there is a maximum number of the most intense links, and between the terms of different clusters there was a minimum possible number of connections and the intensity of the links was the smallest. The connection between terms is reflected both by lines, and in their mutual positioning. The location of terms on the map is the result of a solution of the optimization task, which takes into account the strength of the connections between nodes and their belonging to particular clusters. The analysis of clusters—statistically identified topics—allows us to identify key
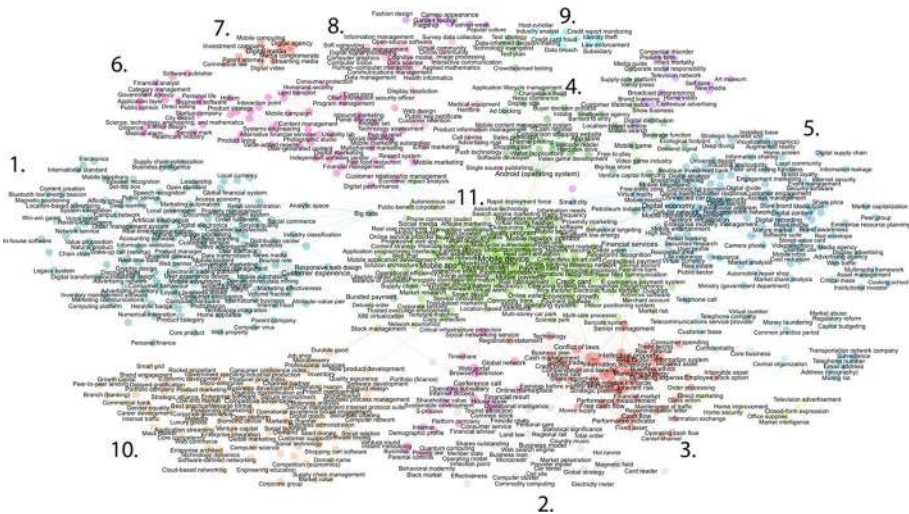
Fig. 3 Semantic graph of global trends in m-commerce industry. *Source*: HSE ISSEK iFORA. (Color figure online)

trends and promising areas of development for m-commerce industry in order to determine their structure.

In total 11 key clusters were identified in the map and described later on as indicators of main trends in m-commerce. These are:

1. **Customer experience**: Customer Relationship Management (CRM) services and instruments, including social commerce
2. **Financial management**, including decentralized financial systems and currencies
3. **Financial analysis and strategy**
4. **Customer-oriented services**
5. **Strategic management**: analysis, market positioning and instruments
6. **Digital and mobile marketing**: strategy, channels, content
7. **Digital media**
8. **Information management**
9. **Data security**
10. **Business planning**: resources and infrastructure management, strategic partnerships
11. **Payment systems**

Each cluster is indicated in Fig. 3 with its corresponding number.

Next sections briefly discuss each cluster and trends identified in association to them. In more details contents of each cluster with the scores are presented in tables in the "Appendix".

1. **Customer experience** cluster includes CRM services and applicable instruments including natural language processing and speech recognition technologies (for more details see Table 1 in the "Appendix").

   One of the key components of the topic under consideration, *social commerce* (score 5), is characterized by Social Experience Design (SXD) concept, which is concerned

with user generated content and self-promoting products and services through auto-mated experience sharing (*marketing automation*—score 5, *real-time communica-tion*—score 4). The developing concept of SXD implies constant share of experience about purchased products and services with the help of mobile devices and social net-works (*customer insight*—score 3). Unification of various social mobile platforms for real-time knowledge exchange and growing amount of total users make SXD relevant not only to the long term planned purchases, but any transaction being committed at the very moment. Thus, social commerce utilises actively the user-generated content using instruments like *natural language processing* (score 5) and *speech recognition* technologies (score 5). In the longer term, the future of m-commerce is expected to lie in the sphere of the content generated not only by users themselves, but also by goods/services, where the Internet of Things (IoT) will play a greater role. Payment processes are already integrated into social networks, where demand and supply sides are coupled and suggest prices and terms of agreement to satisfy both sides. There is a good potential that consumers themselves will use crowdsourcing platforms for direct exchange and delivery of goods and services.

2. **Financial management** and **3. Financial analysis and strategy** are two closely con-nected clusters combining financial instruments for business, including loans, advertis-ing and other types of expenditures (see Table 2 in the "Appendix") as well as more operational management practices in this sphere of m-commerce (see Table 3 in the "Appendix").

Money operations in m-commerce (*cash management*—score 3) include a set of financial services to be delivered over a mobile phone such as mobile wallet, P2P (person-to-person), G2P (government-to-person) and B2P (business-to-person) trans-fers and remittances, payments (including airtime), savings, credit (*microcredit*—score 4, *business loan*—score 2) and insurance among the others.

Among the key enabling technologies for this field are decentralized financial systems, crypto-currencies and universal independent electronic payment methods. Development of m-commerce might contribute to the creation of universal electronic virtual currencies independent from national financial systems. Such currencies would be used for global m-commerce using mobile devices. Many companies and venture funds made investments in the emerging and high growth technology fields enabling mobile payments, digital currency, and mobile and digital financial security.

Emergence of new forms and relations in m-commerce leads to the establishment of a new financial ecosystem for m-commerce support. In such an ecosystem, new money-issuing entities are created such as e-money and crypto-currency. While commerce is becoming more electronic and mobile, mobile payment devices should become an identifier for a person (*card reader*—score 2). This suggests a new level of regula-tion in sales of smartphones and sim cards required to provide security for financial transactions and combat grey economy, tax avoidance, drug trade, and other sorts of corruption.

4. **Customer-oriented services** cluster represents the topic of buyer decision process analysis and shopper marketing (Shankar et al. 2016) strategies development on this basis (see Table 4 in the "Appendix").

Data-informed decision-making in this context is connected with *predictive and personalized mobile commerce*. Machine learning of behaviors and habits (e.g. cus-tomized mobile menus in restaurants), real time financial decision-making (i.e. mobile crediting), mobile marketing (e.g. when point of purchase/service is closer) and multi-dimensional context awareness are among the trends in this category (*buyer decision*

process—score 3). Collection of data about mobile transactions, various queries and browser search history for various customer target groups help to build models and behavioral patterns using algorithms of machine learning (*mobile content management system*—score 5, *mobile search*—score 3, *data-informed decision-making*—score 2). 'Predictive m-commerce' has a potential to make almost all mobile services personalized with unique customer behaviors.

Predictive and personalized m-commerce is also associated with the growth of geo-dependence of mobile apps. Mobile applications using *Location-based Services* (score 5) that take the users' location into account in order to deliver services are becoming widespread. Another associated category is proximity services that inform users when they are within a certain distance to other people, businesses, or places. There are numerous ways to utilise location to provide more relevant information or derive new services. It can be particularly powerful when combined with other user profile information to offer personalized and location sensitive solutions and responses to customers.

5. **Strategic management** cluster includes the topics of market analysis and positioning as well as special management instruments (see Table 5 in the "Appendix").

Mobile technologies and applications have disrupted the traditional business flow and model of retailers. As m-commerce is a highly dynamic field, firms need to develop adaptive capabilities and learn through experimentation for developing innovative responses to changing environment and consumer interests, such as through data-driven management strategies (*behavioral targeting*—score 5, *securities research*—score 3), *customer engagement* (score 4, *engagement marketing*—score 5, *brand loyalty*—score 3, *brand awareness*—score 2), *information sharing* (score 3), *digital supply chain management* (score 1, *branchless banking*—score 5, *electronic bill payment*—score 5, *enterprise resource planning*—score 2) and *mobile identity management* (score 3, *access management*—score 3, *digital identity*—score 2).

C2C business model (*for sale by owner*—score 5) within m-commerce market is bringing many new opportunities for consumers to trade goods and services without the involvement of third parties. However, there are many threats concerning violation of transaction regulations and intellectual rights (*information leakage*—score 2), lack of quality assessment and uncontrolled delivery of prohibited goods.

6. **Digital and mobile marketing** area includes several important concepts and instruments like *social* (score 5) and *multichannel marketing* (score 4), *inbound marketing* (score 5) and *content management* (score 4) as well as *category* (score 2) and *program management* (score 2, for more instruments see Table 6 in the "Appendix").

Mobile shopping combines interactivity and instantaneity as mobile devices in real time provide all possible information over the Internet that can easily lead to abandonment or acceleration of previous shopping plans. Therefore, mobile marketing could effectively strengthen or inhibit shopping goals and increase the number of unplanned purchases with ubiquitous contextual personalized mobile coupons and offers (*inbound marketing*—score 5), etc., because their serendipity and unexpectedness can create a positive effect. At the same time, concerns about privacy and intrusiveness (*consumer protection*—score 2) arise in dynamic tracking of consumer behavior by companies via mobile devices (further elaborated below).

Along with boosting spending, one of the main goals of mobile marketing is to compete for app space on mobile devices. Consequently, m-commerce apps need to be clinging to induce shoppers to continue using them and should be continuously updated and redesigned to remain at the bleeding-edge. Another important aspect

of digital marketing strategy is to keep track of stimulating clients to provide public feedback in social media where they share information, visuals, and video about their shopping experience (*user-generated content*—score 3). To succeed *multichannel marketing* (score 4), marketing professionals have been trying to create more memorable experiences that can be shared with the online community, often with the engagement of opinion leaders, influencers and bloggers.

7. **Digital media** was also automatically identified as an important area of m-commerce field. The cluster #7 in the map indicates the terms of *digital agency* (score 4), *streaming media* (score 4), and *mobile computing* (score 5) along with developing business models and customer relationships modes.

8. Different aspects of **information management** are among the crucial issues in m-commerce (see Table 7 in the "Appendix").

Data related to shoppers' mobile use is voluminous and growing exponentially. Therefore, m-commerce embodies the challenges associated with collecting, managing, and analyzing big data (*information management*—score 3, *data management*—score 3, *knowledge management*—score 2, *digital signal processing*—score 2, *survey data collection*—score 2). Companies invest in searching for instruments (*open-source software*—score 1) to leverage the volume, velocity, variety, and veracity of mobile data and derive value for the company and at the same time enhance data security. In order to assist real time mobile marketing decisions, companies use mobile data in spectrum of domains, ranging from product bundling to dynamic pricing. Survival analysis is one of the instruments enabling prediction of customer switching and signaling the right moment to provide special offers to avert defections (*customer retention*—score 4).

9. Connected with the information management topic, **data security** arose as an important concern in m-commerce (see Table 8 in the "Appendix").

Data security is important both at the technological and policy levels, especially when dealing in C2C business models with the involvement of social networks and personalization of commerce.

Contactless payments, remote payments (*credit card fraud*—score 4), mobile points of sale—all rely on encryption, authentication and authorization services. Elliptic curve cryptography, digital certificates and digital signature wireless LAN security technologies are often used for data transmission and monetary transactions, including m-commerce. Homomorphic encryption brings security by allowing separate packages of data to be transmitted from sending node to receiving node, allowing packages to be then integrated without being decrypted. The use of the same technology for m-commerce, while considering a very big amount of customers' personal data to be constantly processed and passed, is highly anticipated.

Internet of Things (IoT)-related enhancement of processes opens doors for automation of services use without consulting the consumer. Widespread development of m-commerce will induce governments and international regulatory bodies to further regulate the personal data collection via mobile devices (*data breach*—score 3). A serious violation in security would hinder the growth of m-commerce, as well as increase the cost of related operations. Misuse of information and payment details (e.g. credit card details), identity theft and other privacy and security issues would make strong influence on service development as well as policy development. There are also risks related to quality of product or service and m-commerce legislation, transparency and characteristics of safety and security an online vendor offers. All these risks create big mistrust to m-commerce services within society. Solution to these problems comes

from strict legislation. There are many evidences that the development of transnational regulations development would have strong influence on widespread use of e-commerce and m-commerce.

10. **Business planning** topic includes issues related to resource management, infrastructure development, *enterprise integration* (score 2) and *strategic alliances* (score 2) (see Table 9 in the "Appendix").

     Many m-commerce companies struggle to maintain an organization that houses all the required capabilities and resources (*stock management*—score 3, *customer support*—score 2, *new product development*—score 2, *product design*—score 2, *business administration*—score 2, *product marketing*—score 2, *project management*—score 2, *software development*—score 2), creating a need for a network structure to build on. Companies can diversify risk by creating a partnering network (*channel partner*—score 3, *general partnership*—score 2, *open platform*—score 2), because dependencies on one party are likely to make the company vulnerable in highly dynamic environment.

11. **Payment systems** topic is a fundamental one for m-commerce along with the other associated topics in Table 10 in the "Appendix".

     The evolution of B2C and C2C models for retail and other purposes has been strongly influenced by the mobile technologies (*mobile operating system*—score 5) aimed at making transaction process easy and secure. Mobile technologies and *devices* (score 5) are the main instruments of payment and commerce (*mobile payment*—score 5), and have an outstanding potential to shift existing markets. Portable devices have become particularly attractive when versatility, security and simplicity features of payment technologies are considered (*payment processor*—score 5, *host card emulation*—score 4, *payment gateway*—score 3, *location-based commerce*—score 2, *Bluetooth low energy*—score 2). Transactions can be made directly within an application that will bring together all means of payments and will be used as main identifier for individuals' bank accounts (*online banking*—score 2, *multi-factor authentication*—score 2, *fingerprint recognition*—score 2). The service enables customers to make *contactless payments* (score 3, *contactless smart card*—score 2) via mobile devices with a contactless reader or via text-based or personal identification-number-based methods (score 2). Available methods use *Near Field Communication* technology (score 2) that involves short distance communication between the consumer's device, the payment scheme operator, and the retail merchant at the storefront. Infrastructure wise, increase in channel capacity and speed, expansion of 4G, 5G and LTE brings additional benefits of using mobile as the main instrument of payment and commerce even in the context of more traditional online transactions (*social media*—score 4, *online advertising*—score 2, *proximity marketing*—score 2).

Conclusions and discussion

M-commerce is among the key business and technology areas with extremely high market potentials. The recent COVID-19 pandemic has been one of the drivers for a greater demand for m-commerce. In order to meet this demand widespread access to cheap wireless devices with full functionality, faster and stable mobile networks, as well as lower tariffs for mobile communication are required. Increasing involvement of the society in the online activities such as social networks and various mobile applications for personalized

shopping creates many valuable information flows for the exchange between consumers and businesses to offer, rate, compare and commit mobile transactions. Meanwhile, the issues like security, protection of personal data and stability of services should continuously be addressed. These and other trends described in the present study aim at exploring the future possibilities in m-commerce for all stakeholders in the area, as well as anticipating plausible future developments and opportunities which would allow them to develop to a bigger extent and at quicker pace.

More specifically, the present study explored the trends and developments in the m-commerce sector in the 11 identified main areas with discussing future opportunities and threats. Mobile is becoming to be seen as the main tool of payment and C2C commerce and employ technologies like Near-Field-Communication, 3G, 4G and LTE and other payment and data transferring technologies. Connections between identified trends, challenges and technologies are explainable horizontally in relation to the m-commerce application groups, including mobile services, payment, banking, advertising, mobile applications and internet, as well as mobile shopping. Decentralized financial systems and currencies might adjust existing bank systems by making it more transparent and secure, or even replacing it with international independent currencies with no direct control over the money flows. The problem of security and data privacy can be solved using new cryptography approaches like homomorphic encryption, elliptic curve cryptography, anonymization of data used by commerce companies for analytics along with traditional security methods, as well as new data policies induced by governments. M-commerce is becoming more socially relevant through the Social Experience Design concept by exchanging consumers' impressions and suggesting direct trades in social networks. Big data predictive analytics allows analyzing consumer behaviors and suggesting personalized product and service offers to each customer depending on his or her preferences and transaction history. Possible isolation of developing countries from international money and shipment transactions might hinder connection with global m-commerce market, but instead lead to strong national markets.

Development of m-commerce and respective technologies brings a number of opportunities for businesses. New business models developed around mobile services stimulate economies opening new areas for companies to participate. Automation and digitalization of different business processes create more competitive retailing environment that leads to better and more customized services for consumers. Further development of m-commerce also means better access to global products and services bringing better quality of services and less dependency on geographical location of customers.

Despite of the development of m-commerce and respective technologies, as well as adjustments made by governments and business organizations to secure payment systems and personal data of citizens, there remain several threats to be addressed. First, m-commerce strongly depends on trust of customers and willingness to share some of the personal information. Second, high dependence on m-commerce transactions might be risky from the security perspective including threads of intrusions with the aim of remote control of mobile devices and consequently banking accounts. Additional security solutions, as a rule, increase the price of goods and services. Finally, strict national regulations of personal data protection might lead to loss of foreign companies from regional m-commerce markets.

Although the amount of textual data generated worldwide about business and technological developments grows rapidly, there is an absence of systematic horizon scanning exercises in the m-commerce field with its social, technological, economic, policy and cultural impacts. Recognizing this need, our study took m-commerce as a case, and

demonstrated the potential of Big Data augmented horizon scanning for the analysis of business trends and technologies in this domain. The research presented allowed bringing together evidence from numerous media reports, news and other sources of information and structuring it into separate concepts. The semantic graph constructed helped to draw a single web of interconnected technologies, trends, drivers and barriers, hence gave an overview of the whole area of m-commerce in one business intelligence (BI) data mart diagram. This can be considered as a significant contribution to the existing analytical systems and tools drawing on Big Data.

Considering the domain itself, the findings of our research demonstrate the development of m-commerce is dependent on innovations beyond the technology and require a big picture view by taking into account global challenges and trends besides the sectoral dynamics. Furthermore, looking into a longer-term future of a certain domain using Big Data requires focus on the deeper analysis of associated technologies along with a life cycle of its development and interconnections to other R&D areas that impact respective domain or are affected by it. This will give a broader context, for instance for m-commerce innovation, and will provide greater potentials for forecasting future opportunities and challenges in the field.

Based on obtained results it can be suggested that further research can advance the proposed approach further by improving big data text mining proxies of sectoral innovations on the basis of flexible combination of various information sources and advanced technologies including word embedding models like word2vec, sense2vec (modification of word-2vec that takes into account phrases and their senses) and BERT (that also analyses context of words) for finding more latent connections between topics and deeper structure of m-commerce sphere. One of the limitations of current methodology can be due to its over attention on media sources, where it is often the case that so-called hype concepts (like *artificial intelligence*) arise, and products are promoted by particular companies and other interest groups. Combined analysis with additional analytical tools like sentiment analysis or life cycle analysis may be beneficial for more complementary and reliable evaluation of emerging technologies and trends associated to them.

Appendix

Table 1 Main contents of 'Customer experience' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Customer experience	5
Social commerce	5
Mobile advertising	5
Responsive web design	5
Artificial intelligence	5
Big data	5
Location-based advertising	5
Marketing automation	5
User experience design	5
Natural language processing	5
Speech recognition	5
Conversion marketing	4
Consumer behaviour	4
Customer service	4
Loyalty business model	4
Machine learning	4
Mobile content	4
Order management system	4
Customer satisfaction	4
Voice portal	4
Native advertising	4
Electronic trading platform	4
Real-time communication	4
Open-source model	4
Bluetooth low energy beacon	4
Speaker recognition	4
Personal finance	4
Deep learning	4
Mobile purchasing	3
Convergent charging	3
Business intelligence	3
Product management	3
Order fulfillment	3
Supply chain optimization	3
Advertising network	3
Magnetic positioning	3
Network service	3
Inventory management software	3
Internet fraud	3
User interface	3
Content delivery network	3
Branded content	3
Customer insight	3
Computing platform	3
Web banner	3
Data analysis	2

Table 1 (continued)

Terms	Score
Real-time data	2
Accounting software	2
Marketing effectiveness	2
Web property	2
Marketing communications	2
Property management	2
Document management system	2
Domain driven data mining	2
Open standard	2
Advertising inventory	1

Table 2 Main contents of 'Financial management' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Microcredit	4
Web search engine	3
Quantum computing	3
Gross merchandise volume	2
Business loan	2
Card reader	2
Commodity computing	2
Sustainable development	1

Table 3 Main contents of 'Financial analysis and strategy' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Order processing	4
Cash management	3
Performance measurement	3
Strategic management	3
Performance indicator	3
Strategic planning	2

Table 4 Main contents of 'Customer-oriented services' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Comparison shopping website	5
Location-based service	5
Mobile content management system	5
Supply-side platform	4
Customer lifetime value	3
Mobile game	3
Push technology	3
Mobile search	3
Crowdsourced testing	3
Chargeback fraud	3
Buyer decision process	3
Ad blocking	3
Application lifecycle management	3
Cash register	2
Data-informed decision-making	2
Ecological footprint	2

Table 5 Main contents of
‘Strategic management’ cluster
Source: HSE ISSEK iFORA

Terms	Score
Engagement marketing	5
Behavioral targeting	5
Mobile media	5
Branchless banking	5
Virtual reality	5
Augmented reality	5
Virtual currency	5
Electronic bill payment	5
Mobile robot	5
For sale by owner	5
Digital economy	4
Financial services	4
Customer engagement	4
Mobile entertainment	4
Digital security	4
Digital health	4
Mobile identity management	3
Network virtualization	3
Brand loyalty	3
Securities research	3
Asset management	3
Access management	3
Information sharing	3
Mobile broadband	2
Digital strategy	2
Retail banking	2
Business rules engine	2
Digital identity	2
Brand awareness	2
Enterprise resource planning	2
Universal design	2
Store brand	2
Security software	2
Life insurance	2
Event management	2
Internet security	2
Multimedia framework	2
Display advertising	2
Information leakage	2
Stored-value card	2
Product demonstration	1
Market capitalization	1
Digital supply chain	1

Table 6 Main contents of 'Digital and mobile marketing' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Mobile marketing	5
Mobile marketing automation	5
Social marketing	5
Inbound marketing	5
Digital marketing	4
Customer relationship management	4
Multichannel marketing	4
Content management	4
Direct selling	4
User-generated content	3
Event store	3
Product lining	2
Product strategy	2
Category management	2
Service mark	2
Program management	2
Mobile campaign	2
Value-added reseller	2
Consumer protection	2
Alternative financial services	2
Financial management	2
Software publisher	2
Bridge bank	2

Table 7 Main contents of 'Information management' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Email marketing	4
Product information management	4
Computer vision	4
Customer retention	4
Single source publishing	4
Image processing	4
Communications management	3
Information management	3
Human–computer interaction	3
Data science	3
Web design	3
Data management	3
Virtual community	3
Display resolution	2
Digital signal processing	2
Survey data collection	2
Knowledge management	2
Online community	2
Health informatics	2
Advertising mail	2
Interactive communication	2
Open-source software	1

Table 8 Main contents of 'Data security' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Credit card fraud	4
Threat (computer)	3
Credit report monitoring	3
Data breach	3
Hoot-n-holler	2
Industry analyst	1

Table 9 Main contents of 'Business planning' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Stock management	3
Channel partner	3
Business development	3
Competitive advantage	3
New product development	2
Enterprise software	2
Securities offering	2
Professional services	2
Credit (finance)	2
General partnership	2
Venture capital	2
Customer support	2
Global marketing	2
Enterprise integration	2
Strategic alliance	2
Angel investor	2
Product design	2
Seed money	2
Business administration	2
Venture round	2
Message format	2
Digital goods	2
Open platform	2
Consumer confidence index	2
Application streaming	2
Consumer organization	2
Product marketing	2
Software-defined networking	2
Private equity	2
Project management	2
Software development	2
Portfolio company	2
Business process management	2
Social engagement	2
Shopping cart software	2

Table 10 Main contents of 'Payment systems' cluster *Source*: HSE ISSEK iFORA

Terms	Score
Mobile payment	5
Mobile device	5
Mobile app	5
Loyalty program	5
Financial institution	5
Credit card	5
Payment processor	5
Mobile network operator	5
Debit card	5
Mobile operating system	5
Payment card	4
Host card emulation	4
Service provider	4
Social media	4
Issuing bank	3
Payment gateway	3
Payment system	3
Bank account	3
Supply chain	3
User experience	3
Contactless payment	3
Market segmentation	2
Location-based commerce	2
Mobile application development	2
Affiliate marketing	2
Lifestyle center (retail)	2
Proximity marketing	2
Fingerprint recognition	2
Near field communication	2
Mobile security	2
Social network	2
Contactless smart card	2
Operational efficiency	2
Personally identifiable information	2
Media proprietor	2
Online advertising	2
Bluetooth low energy	2
Online banking	2
Value-added service	2
Online service provider	2
Multi-factor authentication	2

References

Bakhtin, P., Saritas, O., Chulok, A., Kuzminov, I., & Timofeev, A. (2017). Trend monitoring for linking science and strategy. *Scientometrics, 111*(3), 2059–2075.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Chiarello, F., Melluso, N., Bonaccorsi, A., & Fantoni, G. (2019, July). A text mining based map of engineering design: Topics and their trajectories over time. In *Proceedings of the design society: International conference on engineering design* (Vol. 1, No. 1, pp. 2765–2774). Cambridge: Cambridge University Press.

Chiou, L., & Tucker, C. E. (2020, April). Social distancing, internet access and inequality. National Bureau of Economic Research working paper no. 26982.

Chuang, J., Gupta, S., Manning, C., & Heer, J. (2013). Topic model diagnostics: Assessing domain relevance via topical alignment. In *Proceedings of the 30th international conference on machine learning (ICML-13)* (pp. 612–620).

Ena, O., Mikova, N., Saritas, O., & Sokolova, A. (2016). A methodology for technology trend monitoring: The case of semantic technologies. *Scientometrics, 108*(3), 1013–1041.

Fong S., & Lai E. (2005). Mobile mini-payment scheme using SMS-Credit. In *Computational science and its applications—ICCSA 2005, PT, Lecture Notes in Computer Science* (pp. 1106–1114).

Gök, A., Waterworth, A., & Shapira, P. (2015). Use of web mining in studying innovation. *Scientometrics, 102*(1), 653–671.

Goldberg, Y., & Levy, O. (2014). word2vec explained: Deriving Mikolov et al.'s negative sampling word-embedding method. arXiv preprint, arXiv:1402.3722.

Groß, M. (2015). Mobile shopping: A classification framework and literature review. *International Journal of Retail & Distribution Management, 43*(3), 221–241.

Harris, W. L., & Wonglimpiyarat, J. (2019). Blockchain platform and future bank competition. *Foresight, 21*(6), 625–639.

Hasanat, M. W., Hoque, A., Shikha, F. A., Anwar, M., Hamid, A. B. A., & Tat, H. H. (2020). The impact of coronavirus (COVID-19) on E-Business in Malaysia. *Asian Journal of Multidisciplinary Studies, 3*(1), 52–58.

Hew, J. J. (2017). Hall of fame for mobile commerce and its applications: A bibliometric evaluation of a decade and a half (2000–2015). *Telematics and Informatics, 34*(1), 43–66.

Hootsuite & We Are Social (2019). Global Digital Report. Retrieved April 7, 2020, https://wearesocial.com/blog/2019/01/digital-2019-global-internet-use-accelerates

Hossain, S. F. A. (2019). Social networking and its role in media entrepreneurship: Evaluating the use of mobile phones in the context of online shopping—A review. *Journal of Media Management and Entrepreneurship, 1*(1), 73–86.

Huang, Y., Zhang, Y., Ma, J., Porter, A., & Wang, X. (2015, August). Tracing technology evolution pathways by combining tech mining and patent citation analysis. In *2015 Portland international conference on management of engineering & technology*.

IDC (2017). As data grows, so does demand for object-based storage. Retrieved February 6, 2019, https://www.ibm.com/cloud-computing/products/storage/object-storage/data-growth-drives-object-storage-demand/index.html

IDC. (2018). The Digitization of the World—From Edge to Core. Retrieved April 7, 2020, https://www.seagate.com/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf

Jamil, S., & Mousumi, F. (2008). Short Messaging Service (SMS) Based m-banking system in context of Bangladesh. In *11th international conference on computer and information technology, 2008. ICCIT 2008* (pp. 599–604).

Ju, M., Nguyen, N. T., Miwa, M., & Ananiadou, S. (2020). An ensemble of neural models for nested adverse drug events and medication extraction with subwords. *Journal of the American Medical Informatics Association, 27*(1), 22–30.

Juventia, S. D., Jones, S. K., Laporte, M. A., Remans, R., Villani, C., & Estrada-Carmona, N. (2020). Text mining national commitments towards agrobiodiversity conservation and use. *Sustainability, 12*(2), 715.

Kayser, V. (2017). Comparing public and scientific discourse in the context of innovation systems. *Technological Forecasting and Social Change, 115,* 348–357.

Kim, S., Park, H., & Lee, J. (2020). Word2vec-based latent semantic analysis (W2V-LSA) for topic modeling: A study on blockchain technology trend analysis. *Expert Systems with Applications, 152,* 113401.

Loesche, D. (2018, March 6). Infographic: Mobile E-Commerce is up and poised for further growth. *Statista Infographics*. https://www.statista.com/chart/13139/estimated-worldwide-mobile-e-commerce-sales/.

Lyytinen, K., & Yoo, Y. (2002). Research commentary: The next wave of nomadic computing. *Information Systems Research, 13*(4), 377.

Miles, I., Saritas, O., & Sokolov, A. (2016). *Foresight for science, technology and innovation*. Berlin: Springer.

Moro, S., Pires, G., Rita, P., & Cortez, P. (2019). A text mining and topic modelling perspective of ethnic marketing research. *Journal of Business Research, 103,* 275–285.

Nair, A. K. S., & Bhattacharyya, S. S. (2019). Is sustainability a motive to buy? An exploratory study in the context of mobile applications channel among young Indian consumers. *Foresight, 21*(2), 177–199.

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2015). Text mining of news-headlines for FOREX market prediction: A Multilayer Dimension Reduction Algorithm with semantics and sentiment. *Expert Systems with Applications, 42*(1), 306–324.

Ngai, E., & Gunasekaran, A. (2007). A review for mobile commerce research and applications. *Decision Support Systems, 43*(1), 3–15.

Okyere, M. A., Forson, R., & Essel-Gaisey, F. (2020). Positive externalities of an epidemic: The case of the Corona Virus (COVID-19) in China. *Journal of Medical Virology*. https://doi.org/10.1002/jmv.25830.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245–251.

Petrock W. (2020, April 6). How COVID-19 is fast-tracking emerging tech: Innovations show renewed promise during pandemic. eMarketer. https://www.emarketer.com/content/how-covid-19-is-fast-tracking-emerging-tech.

Porter, A., & Cunningham, S. (2004). *Tech mining: Exploiting new technologies for competitive advantage*. New York: Wiley.

Pousttchi, K., Tilson, D., Lyytinen, K., & Hufenbach, Y. (2015). Introduction to the special issue on mobile commerce: Mobile commerce research Yesterday, Today, Tomorrow—What remains to be done? *International Journal of Electronic Commerce, 19*(4), 1–20.

Pramanik, M. I., Lau, R. Y., Yue, W. T., Ye, Y., & Li, C. (2017). Big data analytics for security and criminal investigations. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7*(4), e1208.

Rivera, S. J., Minsker, B. S., Work, D. B., & Roth, D. (2014). A text mining framework for advancing sustainability indicators. *Environmental Modelling and Software, 62,* 128–138.

Rong, X. (2014). word2vec parameter learning explained. arXiv preprint, arXiv:1411.2738

Scharl, A., Dickinger, A., & Murphy, J. (2005). Diffusion and success factors of mobile marketing. *Electronic Commerce Research and Applications, 4,* 159–173.

Tang, A. K. (2019). A systematic literature review and analysis on mobile apps in m-commerce: Implications for future research. *Electronic Commerce Research and Applications, 37,* 100885.

The World Bank. (2020) GDP (current US$). Retrieved April 19, 2020 from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD.

Tobback, E., Naudts, H., Daelemans, W., de Fortuny, E. J., & Martens, D. (2018). Belgian economic policy uncertainty index: Improvement through text mining. *International Journal of Forecasting, 34*(2), 355–365.

Trask, A., Michalak, P., & Liu, J. (2015). sense2vec-a fast and accurate method for word sense disambiguation in neural word embeddings. arXiv preprint arXiv:1511.06388.

Vidra, J. (2015). Implementation of a search engine for DeriNet. In *ITAT*. Prague: Charles University in Prague (pp. 100–105).

Wang, Y., Li, Q., Huang, Z., & Li, J. (2019, June). EAN: Event attention network for stock price trend prediction based on sentimental embedding. In *Proceedings of the 10th ACM conference on web science* (pp. 311–320).

World Trade Organization. (2019a, April 2). Global trade growth loses momentum as trade tensions persist. https://www.wto.org/english/news_e/pres19_e/pr837_e.htm

World Trade Organization. (2019b). World Trade Statistical Review 2019. https://www.wto.org/english/res_e/statis_e/wts2019_e/wts19_toc_e.htm.

Yono, K., Izumi, K., Sakaji, H., Shimada, T., & Matsushima, H. (2019, June). Measuring the macroeconomic uncertainty based on the news text by supervised LDA for investor's decision making. In *The International conference on decision economics* (pp. 125-133). Cham: Springer.

Zhang, W. N., Liu, T., Yang, Y., Cao, L., Zhang, Y., & Ji, R. (2014). A topic clustering approach to finding similar questions from large question and answer archives. *PLoS ONE, 9*(3), e71511.

Zhou, H., Yu, H., Hu, R., & Hu, J. (2017). A survey on trends of cross-media topic evolution map. *Knowledge-Based Systems, 124,* 164–175.

Affiliations

Ozcan Saritas[1] · Pavel Bakhtin[1] 🆔 **· Ilya Kuzminov[1] · Elena Khabirova[1]**

Ozcan Saritas
osaritas@hse.ru

Ilya Kuzminov
ikuzminov@hse.ru

Elena Khabirova
etochilina@hse.ru

[1]   Institute for Statistical Studies and Economics of Knowledge (ISSEK), National Research
University Higher School of Economics, Moscow, Russia