International Journal of Information Management 60 (2021) 102331

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Review

From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets

Jose Ramon Saura [a,*], Domingo Ribeiro-Soriano [b], Daniel Palacios-Marqués [c]

[a] Rey Juan Carlos University, Madrid, Spain
[b] IUDESCOOP, Universitat de Valencia, Valencia, Spain
[c] Universitat Politècnica de València, Valencia, Spain

ARTICLE INFO

Keywords:
User-generated data
Data-driven innovation
Users' privacy
Privacy concerns

ABSTRACT

In recent years, strategies focused on data-driven innovation (DDI) have led to the emergence and development of new products and business models in the digital market. However, these advances have given rise to the development of sophisticated strategies for data management, predicting user behavior, or analyzing their actions. Accordingly, the large-scale analysis of user-generated data (UGD) has led to the emergence of user privacy concerns about how companies manage user data. Although there are some studies on data security, privacy protection, and data-driven strategies, a systematic review on the subject that would focus on both UGD and DDI as main concepts is lacking. Therefore, the present study aims to provide a comprehensive understanding of the main challenges related to user privacy that affect DDI. The methodology used in the present study unfolds in the following three phases; (i) a systematic literature review (SLR); (ii) in-depth interviews framed in the perspectives of UGD and DDI on user privacy concerns, and finally, (iii) topic-modeling using a Latent Dirichlet allocation (LDA) model to extract insights related to the object of study. Based on the results, we identify 14 topics related to the study of DDI and UGD strategies. In addition, 14 future research questions and 7 research propositions are presented that should be consider for the study of UGD, DDI and user privacy in digital markets. The paper concludes with an important discussion regarding the role of user privacy in DDI in digital markets.

1. Introduction

In the beginning of the 21st century, the development of data-centric strategies has changed the paradigm and existing business models (Lies, 2019). Strategies focused on data-driven innovation (DDI) have led to the emergence and development of both new products, business models and opportunities in the digital ecosystem (Akter et al., 2019; Bouncken, Kraus, & Roig-Tierno, 2019; García-Cabrera, García-Soto, & Olivares-Mesa, 2019). The digital ecosystem consists of digital markets where the information generated as a result of user actions is stored in the form of data. These data can then be analyzed in order to find patterns and trends (from de Camargo Fiorini, Seles, Jabbour, Mariano, & de Sousa Jabbour, 2018).

Likewise, recent advances in both information and data sciences have led to the emergence of sophisticated strategies in companies for data management, the ability to make various predictions by applying artificial intelligence, and the application of concepts related to data automation, marketing intelligence or business intelligence (Hargittai, 2010). Digital markets have come to be understood as social networks, large marketplaces, and any digital platform that brings together traffic from individual users that can be identified in forms of online communities (Liu & Lai, 2020; Öberg & Alexander, 2019).

Users are usually structured in digital communities where individuals share their interests and concerns about products and services, as well as communicate about companies, thereby fostering increased engagement (Allen & Shoard, 2005; Ricciardi, Zardini, & Rossignoli, 2018).

The data that emerge as a result of these user actions is referred to as user-generated data (UGD). UGD includes all forms of information and data that users generate individually as a result of interacting with the elements that make up any digital market (actions, experiences, feelings, comments, reviews, and so forth) (Saura, 2020).

Overall, data analysis strategies have been extensively studied in the literature (Stieglitz, Mirbabaie, Ross, & Neuberger, 2018; Vanhala et al., 2020; Yu, Zhang, Lin, & Wu, 2019). These and other relevant studies define the techniques used to collect, structure, analyze, and interpret

* Corresponding author.
E-mail addresses: joseramon.saura@urjc.es (J.R. Saura), domingo.ribeiro@uv.es (D. Ribeiro-Soriano), dapamar@doe.upv.es (D. Palacios-Marqués).

https://doi.org/10.1016/j.ijinfomgt.2021.102331
Received 24 November 2020; Received in revised form 7 February 2021; Accepted 8 February 2021
Available online 19 February 2021

large amounts of data (Ferreira & Teixeira, 2019). All these approaches are contextualized under the concept of Big Data Analytics (BDA).

In the framework of BDA and UGD, companies are developing strategies focused on increasing their profitability in the digital markets. Yet, these strategies can lead to concerns regarding user privacy (e.g., Arya et al., 2019; Zuboff, 2015). This occurs because, rather than relying on functionality, useful information architecture, or user experience while maintaining an ethical design, companies prioritize their economic objectives (Bandara, Fernando, & Akter, 2020; Zuboff, 2019).

Since users may not be aware of being manipulated in digital markets through advertising, design of information architectures, or prediction of behavior, several previous studies have highlighted the importance of concepts such as surveillance capitalism or ethical design in social networks (Zuboff, 2019). These approaches are usually designed by DDI that companies test in their digital ecosystems (Pangrazio & Selwyn, 2019).

In surveillance capitalism, privacy of users and their data in digital environments must prevail over the economic interests of large technological multinationals and governments (Roberts, 2015). The ethical design in digital environments should be a priority for companies. In this respect, Hawi and Samaha (2017) demonstrated that companies can use data to benefit from users economically (González, 2017).

Companies use both DDI and BDA to innovate in their analytical development strategies, in an attempt to identify patterns in large databases generated by user actions and to improve their decision making. With these BDA analyses, companies modify the information structure of their sites, thereby increasing the possibility of achieving engagement as a key part of the interaction and data generation between users and the company (Isaak & Hanna, 2018).

Many previous studies have highlighted the link between new products and services focused on DDI and user privacy (Zuboff, 2015; Arya et al., 2019). Sometimes users are not aware that, as a result of their actions online (IoT, mobile devices, social media profiles, mobile applications, etc.), they are generating data that can be later used by companies to gain economic benefit (Paine, Reips, Stieger, Joinson, & Buchanan, 2007). If these datasets are studied using Artificial Intelligence, machine learning, deep learning, or BDA (Kar & Dwivedi, 2020), it will be possible to considerably better predict user actions, which would also enhance the risk of user privacy violations in digital ecosystems (Gutierrez, O'Leary, Rana, Dwivedi, & Calle, 2019). Despite the growing concerns about user security and privacy, more and more data are generated, and users continue to share information, create content, and spread their messages and opinions on the Internet (Baird & Fisher, 2005). In addition, the emergence of DDI models and strategies to track user data (predictive algorithms, machine-learning, cookies, beacons, etc.) has led to the emergence of databases that, instead of collecting content, gather behavioral data of users in digital markets. This type of content is known as User-Generated Behavior (UGB) (Netzer, Tenenboim-Weinblatt, & Shifman, 2014; Vanhala et al., 2020).

In this context, the present study aims to investigate the link between the generation of new products and services focused on DDI and UGD as sources of data, as well as to explore the consequences these strategies may have for user privacy. Moreover, we also explore how UGD can be used by DDI to generate safe and consistent strategies that do not violate user privacy in digital markets, which fills a gap in the literature by the analysis of user privacy from the DDI and UGD perspectives. The main research question addressed in the present study is as follows: *What are the challenges of DDI models in digital markets in the context of increasing user privacy concerns?*

To answer this research question, we aim to accomplish the following objectives:

- To identify definitional perspectives of user privacy in DDI from the UGD theoretical perspective
- To explore the types of DDI approaches to preserve user privacy in digital markets

Table 1
Theories of user-generated data production in digital markets.

	Description	Authors
Critical-mass theory	This theory posits that, provided there is a sufficiently high number of supporters of an idea, technology, innovation, or social system, the adoption of this idea, technology, etc. will be self-sustaining and will cause its growth.	Prasarnphanich and Wagner (2009) Peng (2010) Sledgianowski and Kulviwat (2009)
Information overload theory	According to this theory, when a large amount of input into a digital system exceeds its capacity for data processing, the information overload will lead to worse decisions, as the cognitive processing capacity is limited.	Kaufhold et al. (2020) Ndumu (2019) Saxena and Lamest (2018) Allen and Shoard (2005)
Common-ground theory	This theory argued that the overlap between different opinions and positions on a subject may lead to disagreements. In digital markets, this phenomenon leads to the appearance of large amounts of data with segmented feelings and personalities.	Keller et al. (2017) Westerman et al. (2014) Schoen et al. (2013) Wohn, Lee, Sung, and Bjornrud (2010)

Source: The authors.

- To create knowledge about the use of UGD in DDI preserving user privacy
- To provide future guidelines to track the challenges of DDI with regard to privacy

In terms of methodology, the approach adopted in the present study unfolds in the following three steps. First, we undertake a systematic literature review (SLR). Second, based on its findings, we conduct in-depth interviews with leading professionals of the IT industry. Thirdly and finally, we employ a Latent Dirichlet allocation (LDA) model and a textual analysis (TA) to extract insights relative to the object of study using keyness as a statistical measure that values the log-likelihood score of the results. Based on the results, we identify a total of 14 topics related to the study of DDI and UGD strategies. Furthermore, 14 future research questions and 7 research propositions are identified that must be taken into account in future analysis strategies focused on the use of UGD, DDI, considering the user privacy in digital markets. The paper concludes with an important discussion regarding the role of user privacy in DDI in digital markets.

The remainder of this paper is structured as follows. In Section 2, we present the theoretical framework of the study. Section 3 discusses the methodology. The results are reported in Section 4. In Section 5, we provide a discussion of important theoretical contributions that our results offer for the analysis of DDI in digital markets with respect to privacy of the UGD, as well as discuss future research agenda regarding the role of user privacy in DDI in digital markets. Conclusions are drawn in Section 6.

2. Theoretical framework

In order to understand the theoretical framework that encompasses the development of strategies focused on DDI and UGD, in this section, we review the main theories on the production of UGD in digital markets (Kaufhold, Rupp, Reuter, & Habdank, 2020; Keller, Schoch, Stier, & Yang, 2017; Prasarnphanich & Wagner, 2009;), the characteristics of the UGD in DDI strategies (Karegar, Pettersson, & Fischer-Hübner, 2020; Saura, 2020), the types of intentionally vs. non-intentionally generated consumer data (Schoen et al., 2013; Vanhala et al., 2020), and, finally, the types of trust in UGD in digital markets (Hajli, 2014; Panahi, Watson, & Partridge, 2016), since they encompass both the analysis data-centric approaches, such as trust building, and the study of user behavior in

Table 2
User-generated data characteristics for data-driven innovation.

Characteristics	Description	Key points	Authors
Topic and purpose	Content categories and objectives in social networks	Interest and relevance	Stieglitz et al. (2018) Lozano, Schreiber, and Brynielsson (2017) Törnberg and Törnberg (2016)
Member characteristics	Profile type, user, and customization	Personal information	Hargittai (2010) Su and Contractor (2011) Chen, Vorvoreanu, and Madhavan (2014)
Trust and security	Trust and security in the digital market	Level of trust / Perception of privacy	Cheng, Fu, and de Vreede (2017) Hansen, Saridakis, and Benson (2018) Sembada and Koay (2019)
Usability / UX	Usability of data based on the ecosystem where they are generated	Level of user experience	Tenkanen et al. (2017) Baird and Fisher (2005)
Group/ Community size	Size of the user community around which the data are generated	Power to bring about change	Martinez and Walton (2014) Roberts (2015)
Time factor	Time horizon of subscription or use of a product that generates UGD	Durability	Saura (2020) Stieglitz et al. (2018)
Membership life cycle (cookies)	UGD related to the development of plans and subscription program	Loyalty generates more confidence	Sembada and Koay (2019) Lies (2019)

Source: The authors.

digital markets (Kar & Dwivedi, 2020).

Of note, the public and free access to large amounts of data has provided the companies an opportunity to implement massive advertising campaigns, perform active listening in social networks, as well as offered them an array of commercial opportunities (Sembada & Koay, 2019). This easy access to large amounts of data has also driven companies to increase their data collection and compilation capacities in order to be used to improve managerial decision making (Saxena & Lamest, 2018).

In order to understand how data can help companies to create DDI models and make decisions, we should first consider how data are produced in digital markets (Pangrazio & Selwyn, 2019). From the perspective of UGD analysis, there are different theories about data production in digital ecosystems (see Table 1) and user motivations. These theories support the generation of UGD in digital markets, which is the main source of data for companies.

As indicated above, UGD emerge from intentional user publications and are a consequence of user actions in digital environments (Karegar et al., 2020). The analysis of these data—including user experiences, time of use, or personality types—allow companies to better understand user intentions and predict their behavior. Overall, UGD are derived from (i) information exchange, (ii) common activities, (iii) ideology/religion or (iv) purchase transactions (Karegar et al., 2020; Saura, 2020).

Therefore, the UGD has brought about the opportunity to access a multitude of data sources previously unavailable to companies (see Table 2). These data sources can serve as the basis for the generation of new behavior prediction models, classifying target audiences,

Table 3
Intentionally vs. non-intentionally generated consumer data (UGC and UGB).

		User-generated data (UGD)	
User Data points	Possible sources	User-generated content (UGC) Intentionally generated data	User-generated behavior (UGB) Non-intentionally generated data
Geographic	Apps, mobiles devices	√	√
Categorization/ topical	Social media profiles interests		√
Demographic (Age/Gender)	Profile preferences	√	
Marital status	Search terms, searched content, interactions	√	
Lifestyle	Content consumed, subscriptions		√
Psychographics	user activity, Content consumed; content created	√	√
Household income	Type of products bought, content consumed, subscriptions		√
Family size	Family memberships, number of devices per IP		√
Interests	Content created, users/influencers followed,		√
Opinions	digital platforms, markets and social media profiles	√	
Browsing history	Search engines	√	
Purchase history	Digital marketplaces, e-commerce profiles		√
Time in social media	Social media profiles and apps		√
Ad interactions	Digital markets, social networks, emails		√
Types of media consumed	Digital markets, social networks		√
Search terms used	Search engines, websites search engines	√	
Bank company	Apps downloaded		√
Sports	Apps downloaded, interests in social media, followed users		√
Nearby connected devices (Location)	Mobile devices, Wi-Fi access, location and connectivity		√
Music	Users profiles followed, type of apps downloaded, subscriptions	√	
Education level	Content consumed, institutions followed, professional social networks		√
Health information	Medical apps, e-health services installed in devices		√
Ideology	Comments, users followed,	√	√
Photos	Places visited, users' social connections		√
Text messages	Subscriptions confirmations	√	

(continued on next page)

Table 3 (*continued*)

User Data points	Possible sources	User-generated data (UGD) User-generated content (UGC) Intentionally generated data	User-generated behavior (UGB) Non-intentionally generated data
	messages, transactional data		
Calls	Frequency numbers called, number of closed contacts	√	
Dialog and social interaction	Content consumed, interests, personality	√	
Video-tracking	Personality, type of content consumed		√
Voice	Personality, social lifestyle	√	√
Facial recognition	Personality, mean personal traits		√

Source: The authors.

retargeting campaigns, digital segmentation, and so forth.

Among the technologies used to collect and analyze UGD are novel computer science methodologies such as BDD, data mining, knowledge discovery, machine learning, artificial intelligence (AI), among many others (Saura, 2020).

Collectively, these technologies allow companies to design and implement new strategies based on innovation to gain new added value and competitive advantage or to perform the analysis of new markets (Imran-Daud, Sánchez, & Viejo, 2016). In order to become part of this new game system, companies have adopted new and innovative strategies for the collection, analysis and processing of these datasets (Bandara, Fernando, & Akter, 2020; Judson, Devasagayam, & Buff, 2012). The final objective is to trend the behavior of users in digital markets and then use this knowledge to segment advertising so that to increase profits and profitability (Pratesi, Gabrielli, Cintia, Monreale, & Giannotti, 2020).

Accordingly, companies apply DDI models to better understand user behavior and develop strategies focused on information management and decision making (Prince, 2018). To encourage remarketing strategies (i.e., making social ads pursue users through different devices) or retargeting (i.e., personalizing content based on cookies), companies have devised new ways of collecting data, some of which have brought about the issue of user privacy (Roberts, 2015).

In general, privacy concerns are determined by how companies generate knowledge from the data that users produce in digital markets (Schoen et al., 2013). In this respect, an important concept is that of user data points, i.e., the contacts that users make with applications, devices, and technologies by providing personal information (Vanhala et al., 2020).

The possible sources of such user data are the questions that explicit ask users to provide these details; alternatively, these data can be inferred from user actions online. Accordingly, UGD can be categorized into (i) user-generated content (UGC, i.e., information that users know they are creating publicly) and (ii) user-generated behavior (UGB, i.e., information that is generated as a result of user actions) (see Table 3).

The ease of collecting such personal data from users has led researchers to explore what kind of confidence users have in digital market technologies and ecosystems (Westerman, Spence, & Van Der Heide, 2014; Yu et al., 2019; Zhou, Wu, Wei, & Dong, 2019).

Therefore, UGD allows us to understand the types of trust that users have in these ecosystems, as, by understanding these, companies adapt their strategies for the development of different types of DDI (see Table 4).

By understanding the types of trust that users can develop in digital markets, companies can adapt DDI models to extract, analyze, and monetize user data based on developing strategies that increase user trust and, therefore, the amount of content that users publish in digital markets (Sembada & Koay, 2019; Sledgianowski & Kulviwat, 2009).

3. Methodology

3.1. *Systematic review of literature*

Following Sarkis, Zhu, and Lai (2011), Akter and Wamba (2016), de Camargo Fiorini et al. (2018) and Akter et al. (2019), in the present study, we develop a systematic review of literature to analyze the main academic contributions related to the topic of user privacy and data-driven innovations. As argued by Stieglitz et al. (2018), a literature review is an effective methodology to identify emerging issues that could potentially benefit theoretical foundations related to the object of study—in our case, privacy concerns related to the use of DDI products

Table 4
Types of trust in user-generated data in digital markets.

Type of trust	Description	Key elements	Authors
Interpersonal trust	User perceptions of actions of other people that would harm them. An individual user is willing to accept vulnerability or risk based on expectations regarding another person's behavior.	User perception User interest User vulnerability User expectations	Panahi et al. (2016) Dutta and Bhat (2016) Martinelli Watanuki and de Oliveira Moraes (2019)
System trust	User perceived security or reliance on both the platform system and the community they belong to.	Perceived security in the system User reliance Perceived security in the community of users	Hajli (2014) Wu et al. (2016) Ceron (2015)
Dispositional trust	User confidence in others, independently of context or third-party users.	General attitude Trustworthiness toward trust Independent trust	Szymczak, Kücükbalaban, Lemanski, Knuth, and Schmidt (2016) Utz and Krämer (2009) McKnight, Kacmar, and Choudhury (2004)
Perceived competence	Determined by factors such as secure payments, data privacy, data protection, system responsibility toward data, transparency, adequate access, third-party data sharing, etc.	Perceived data security Perceived data privacy Trust in the system transparency	Areepattamannil and Santos (2019) Tsvere, Swamy, and Nyaruwata (2013) Hajli (2014)
Perceived goodwill	Good intentions and trustworthiness of community members to develop interpersonal trust. The higher the perceived goodwill, the more content users will generate, the more personal data will be shared, and the more trust they will have in the community they belong to.	System good intentions and trustworthiness Perceived goodwill	Spence, Lachlan, Westerman, and Spates (2013) Omilion-Hodges and Rodriguez (2014) Judson et al. (2012)

Source: The authors.

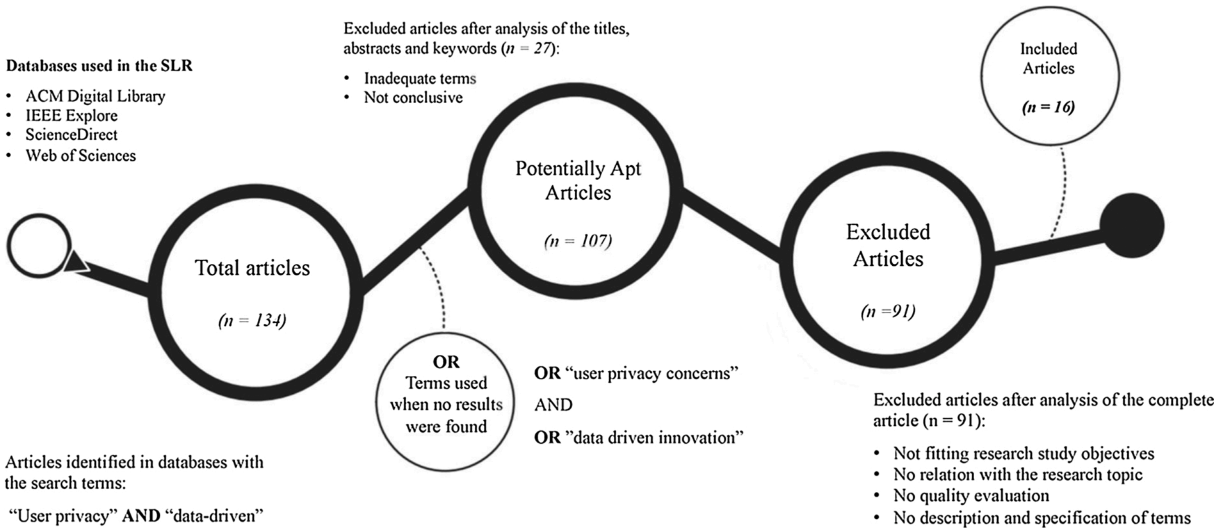

Fig. 1. The SLR process.
Source: Adapted from Saura (2020)

Table 5
Risk bias assessment of the studies included in the "theoretical contributions" category.

Authors	SD	RSG	BOA	WDO	IEC	RAE
Karegar et al. (2020)	+	?	?	+	+	+
Liu and Terzi (2010)	+	+	?	+	+	?
Malgieri and Custers (2018)	?	–	–	–	+	?
Prince (2018)	+	?	?	+	+	+
Tahir et al. (2020)	?	–	–	–	+	–
Yang, Xiong, and Ren (2020)	?	–	–	–	+	–
Yeon Cho et al. (2018)	+	–	–	+	+	+
Yu et al. (2019)	+	+	+	+	+	+

Yes= + No = - Doubtful =?
Source: The authors.

Table 6
Risk bias assessment of the studies included in the "data-driven models" category.

Authors	SD	RSG	BOA	WDO	IEC	RAE
Cheung and She (2016)	+	+	?	–	+	–
He et al. (2020)	+	+	?	–	+	+
Imran-Daud et al. (2016)	+	?	?	–	+	?
Pratesi et al. (2020)	+	?	?	–	+	+
Qi et al. (2020)	+	?	–	–	+	–
Qian et al. (2016)	+	–	?	–	+	–
Zhong et al. (2020)	+	?	+	?	+	–
Zhou et al. (2019)	+	+	+	+	+	–

Yes= + No = - Doubtful =?
Source: The authors.

in digital markets.

For our SLR, following Bem (1995), we first reviewed the theoretical and academic foundations of previous research on UGD and DDI. Next, we identified the main topics discussed within these two areas of research. Finally, based on the two steps mentioned above, we decided on the keywords and their combinations to be used in subsequent database search (Kraus, Breier, & Dasí-Rodríguez, 2020).

Following Stieglitz et al. (2018), our SLR was based on the papers in reputed the academic databases: ACM Digital Library, IEEE Explore, ScienceDirect, and Web of Sciences (WOS). We also considered searching the AIS Electronic Library database; however, the results of searching this database yielded only proceedings (rather than research articles), so this database was not included in our review. In this decision, we followed suggested by Stieglitz et al. (2018) and Saura (2020).

The terms used in the SLR were "User privacy" OR "user privacy concerns" AND "data-driven" OR "data-driven innovation". We used the term "user privacy concerns" when the search of the terms "User privacy" AND "data-driven" did not yield the expected results. The searches were performed on October 12–14, 2020. We focused on titles, abstracts, and keywords to identify relevant contributions. The total number of articles obtained in the search was 134, of which 16 met the inclusion criteria (see Fig. 1).

The number of studies found in the databases was as follows: ACM Digital Library, 35 results, of which 3 met the inclusion criteria, IEEE Explore 46/2, ScienceDirect 46/6, WOS 7/5. As mentioned previously, the total number of results was 134 articles, of which 16 were selected as relevant.

The final step in the review process was to conduct in-depth reading of the identified papers to identify the main contributions and gaps for future research. The 16 articles in the final dataset were analyzed in depth in relation to the theories and definitions identified in the theoretical framework process.

Consequently, the 16 articles were classified into the following two groups: (i) theoretical contributions and (ii) data-driven models (see Tables 5 and 6). In (i), we classified papers that made theoretical contributions to research on user privacy and data-driven innovations. In (ii), we classified the studies that contributed solutions to user privacy in digital markets with the use of data-driven models.

Additionally, in order to ensure accuracy and precision of the reviewed articles (Kiss, Williams, & Houghton, 2013), we performed an assessment of risk bias in both groups of studies taking into account study design (SD), random sequence generation (RSG), blinding of outcome assessment (BOA), withdraw and drop out (WDO), inclusion-exclusion criteria (IEC) and reporting adverse events (RAE) (Table 7).

3.2. In-depth interviews

Next, in order to address our research question and acquire additional knowledge regarding the challenges related to user privacy in data innovation, we conducted a series of qualitative interviews with leading IT professionals (MacDougall & Fudge, 2001). In doing so, our aim was not to achieve statistical generation or significance, but rather to gain an in-depth understanding of the structure of the studied phenomenon (Orlikowski & Baroudi, 1991; Roberts, 2015).

We conducted a total of 11 interviews on data privacy and data

Table 7
provides further detail on the 16 identified articles (the authors, journal, category, classification, main definitions, and contributions to GDA and DDI). Results of Systematic Literature Review.

Authors	Journal	Category	Theoretical contributions	Data-driven models	Purpose	Main concepts analyzed
Cheung and She (2016)	*Transactions on Multimedia Computing, Communications, and Applications*	Multimedia Information Systems		•	To study the privacy issues in online social networks from the individual users' viewpoint	Real-world data, sensitivity, visibility of information and data management
He et al. (2020)	*ACM Transactions on Interactive Intelligent Systems*	Human Computing Interaction		•	To propose a data-driven approach to design privacy-setting interfaces for users in household IoT industry	Developing privacy profiles, privacy default settings, user's privacy preferences
Imran-Daud et al. (2016)	*Computer Communications*	Social and Information Networks		•	To automatically detect sensitive information according to the privacy requirements of the publisher of data.	Privacy-driven access control, content-driven protection of user publications, textual messages, content
Karegar et al. (2020)	*ACM Transactions on Privacy and Security*	Security and Privacy	•		To investigate how interactions that engage users with consent forms differ in terms of their effectiveness, efficiency, user satisfaction and privacy concerns	User privacy and engagement, user attention and satisfaction, types of interactions
Liu and Terzi (2010)	*ACM Transactions on Knowledge Discovery from Data*	Privacy, Social Network Services	•		To approach the privacy issues in online social networks from the individual users' viewpoint proposing a framework to compute the privacy score of a user	Privacy scores, users in online social networks, privacy issues, privacy risk
Malgieri and Custers (2018)	*Computer Law & Security Review*	Computer Technology	•		To analyze whether consumers/users should have a right to know the value of their personal data.	Data-driven economy, pricing privacy, user personal data.
Pratesi et al. (2020)	*Data & Knowledge Engineering*	Data Design, Data Base Tools		•	To analyze privacy issues related to the sharing of user profiles, derived from mobile phone data.	Privacy risk assessment, risk-users, quality of user profiles, user classification of privacy
Prince (2018)	*International Journal of Human-Computer Studies*	Human Computer Interaction	•		To assess the factors that affect web users' predisposition to exert control over personal data flows that targets online users and their privacy	Privacy controls over data flows, concerns over information privacy, individuals' privacy empowerment
Qi et al. (2020)	*Information Sciences*	Information Science		•	To propose a data-driven service recommendation with privacy-preservation.	Collaborative filtering, context-aware, temporal information of service invocations, privacy, decision-making
Qian et al. (2016)	*IEEE Transactions on Computers*	Data Privacy, Computer Science		•	To propose a data-driven analysis which encrypts users' sensitive data to prevent privacy disclosure and to evaluate a real online behavior dataset	Online user behavior data, behavior data, privacy protection, privacy disclosure
Tahir et al. (2020)	*IEEE Access*	Computer Architecture	•		To review the state-of-art application of Blockchain in 5 G network and explore how it can facilitate enabling technologies to use user data.	5 G technology, connectivity, users' perceptions, new technologies testing
Yang et al. (2020)	*IEEE Access*	Cloud Computing, Data Privacy	•		To review the literatures on data security and privacy issues, data encryption technology, and applicable countermeasures in the cloud storage system	Cloud storage, user's data security, user's privacy protection, information disclosure, privacy disclosure
Yeon Cho et al. (2018)	*KSII Transactions on Internet and Information Systems*	Information Systems	•		To investigate factors considered in privacy calculus of fitness devices and verify differences among users	Information privacy, collect sensitive data, privacy concerns
Yu et al. (2019)	*Industrial Marketing Management*	Information Management	•		To construct a conceptual model based on the effects of consumer perceptions of personalized online ads.	Consumer perceptions to ads on the click-through intention, privacy concerns, social content, trust
Zhong et al. (2020)	*Computer Communications*	Computer privacy & Communication Networks		•	To propose a multi-dimensional quality ensemble-driven recommendation approach to make privacy-preserving recommendations	Privacy-preservation, service recommendations, quality ensemble-driven recommendation
Zhou et al. (2019)	*IEEE Journal on Selected Areas in Communications*	Privacy & Data models		•	To analyze subjective privacy-aware evaluation issues of users using a data-driven model	User privacy-aware preferences, observable user data, privacy issues of social network behaviors

Source: the authors.

innovation with the professionals from 9 companies. The informants were from medium and large companies with extensive experience in developing strategies in digital markets (see Table 8). Our interviews were semi-structured and included open-ended questions (see Annex 1).

The main reason for using open-ended questions was to try to address a wider range of experiences (Dhillon & Torkzadeh, 2006). Each interview lasted approximately one hour and was conducted between October 21, 2020 and November 15, 2020.

The interview data were then transcribed and coded using exploratory data-based techniques (Bacq, Janssen, & Noël, 2019;

Table 8
Interviewees by sector, company and professional.

Informant	Sector	Company size	Role of informant	Core duties	Organization Type
A	Telecom	Medium	Senior CTO	Marketing and Communications	Multinational
B	IT	Medium	(i)Senior Computer Scientist, (ii)Digital Marketing Manager	Data Sciences, Digital Marketing	Private
C	Vehicles Industry	Large	Senior Consultant	Communications	Private
D	Marketing	Medium	Digital Marketing Manager	Digital Marketing	Private
E	Software Deve.	Medium	Quality Manager	CRM and Development	Private
F	e-Health	Large	Communication Manager	Media Communications	Private
G	Communi.	Large	Big Data Manager	Marketing	Private
H	Education	Large	Senior CEO	General Management	Private
I	IT	Large	(i)User Experience Manager, (ii)SEO Manager	Media Communications and Design	Multinational

Source: The authors.

Table 9
Demographic characteristics of the interviewees.

Demographic Characteristic	Sub-Level	Count	(%)
Gender	Male	8	66.6
	Female	4	33.3
	Other	–	–
Profession	CTO	1	8.3
	Computer Scientist	1	16.6
	Digital Marketing Manager	2	8.3
	Senior Consultant	1	8.3
	Quality Manager Communication Manager	1	8.3
	Big Data Manager	1	8.3
	CEO	1	8.3
	User Experience Manager	1	8.3
	SEO Manager	1	8.3
		1	8.3
Education	Postgraduate	9	75
	PhD	3	25
Age	26–35	4	33.3
	36–45	5	41.6
	46–56	2	16.6
	> 55	1	8.3

Source: The authors.

Table 10
Data sources in the LDA application.

Characteristics	Ye et al. (2011)	Lee and Bradlow (2011)	Büschken and Allenby (2016)	Hao et al. (2017)	Present Study
Online rating	√	–	√	√	
Comments		√	√	√	
LDA	√		√	√	
Social interactions			√	√	
Interviews					√
Topic frequency		√	√		√

Source: The authors.

Cooke-Davies & Arzymanow, 2003). The demographic characteristics of the informants are summarized in Table 9 based on their professions.

3.3. Data mining: topic-modeling and textual analysis

In recent years, data mining techniques, such as modeling and textual analysis, have come to be extensively used in the literature (Amin et al., 2019; Jimenez-Marquez, Gonzalez-Carrasco, Lopez-Cuadrado, & Ruiz-Mezcua, 2019). In the present study, we used two techniques of data-based approaches. First, a model based on mathematical and probabilistic functions, known as LDA, was applied to analyze the content of the interviews (Blei, Ng, Jordan, & Lafferty, 2003; Pritchard, Stephens, & Donnelly, 2000). The originality of our study lies in that, while LDA has previously been used to analyze the content extracted from social networks and digital markers, in the present study, we used this technique to analyze our interview data by following Krippendorff (2013) content analysis considerations.

Table 11
Characteristics of textual analysis.

Characteristics	Boyd et al. (2010)	Rosa et al. (2015)	Jiang et al. (2016)	Ramirez-Andreotta et al. (2016)	Present Study
Classification into nodes	√	√	√	√	√
Categorization			√		
Word count	√			√	√
Keywords	√	√		√	

Source: The authors.

In general, the LDA model identifies keywords within the analyzed documents and proposes a distribution of themes in a randomly identified sample. Specifically, the model shows the ten most relevant words in the database and based on the results, the researchers can propose different themes. These themes will be the topics that make up the analyzed database (see Table 10 for a review of similar studies). In the present study, this approach was performed using Python software LDA 1.0.5.

In order to ensure that the analyzed topics are relevant, the concept of keyness, also known as the strength of the link, has previous been applied. Keyness is a statistical measure that values the log-likelihood score (Rayson & Garside, 2000; Reyes-Menendez, Saura, & Stephen, 2020). This metric provides statistical meaning and makes it possible to determine differences between two corpora. Specifically, the log-likelihood score of 3.8 or higher was reported to be statistically significant at $p < 0.05$ (Minhas & Hussain, 2014; Reyes-Menendez et al., 2020). Therefore, in the present study, the conversations from the interviews were put into different in-puts phrases and text documents that were considered as sub-corpus and were then compared with the original corpus composed of the full texts collected from the in-depth interviews.

For the set of identified topics, the statistical significance was $p < 0.05$. According to Iyengar, Sood, and Lelkes (2012) and Reyes-Menendez et al. (2020), this allows for measuring log-likehood to determine the importance of the identified topics in the overall analyzed content.

Secondly, we also performed textual analysis with data-mining techniques (Krippendorff, 2013). To this end, different phrases and concepts were grouped in nodes, or text groupings that discussed similar issues. Each node had different variables to measure to evaluate the relevance of the words and concepts that composed it. Specifically, we measured the frequency and repetition of the words within the database, and then the total weight of those word groupings in nodes within the database was measured (Hilal & Alabri, 2013) (see Table 11 for a review of similar studies).

Table 12
Topics identified in interviews.

Topics	Topic description	Keyness	*p*-value
User privacy preferences	Users' preferences regarding their privacy	776.72	0.039
User engagement	Analysis of the type of actions and user engagement	497.80	0.027
Privacy risk	Risks relating to the privacy of user data	417.02	0.024
Data-driven economy	Economics based on data-driven approaches	390.94	00.23
User behavior	Study of user behavior in digital markets	390.03	0.023
Information management	Decisions taken by company management of	379.58	0.021
Decision making	Influence of data-driven innovation and improved decision making	305.11	0.014
User perceptions	User perceptions of security and risk on filtering personal user data	269.08	0.011
Driven content	Actions, techniques, and models focused on data-driven content	164.10	0.008
Social ads	Influence of data-driven models on social ads	135.75	0.006
Sensitive data	Access to sensitive data to study online user behaviors	135.32	0.006

Source: The authors.

Table 13
Grouped keyword nodes.

Keywords	Count	WP
Data driven-innovation, data-driven economy, data-driven models, data points, data-driven behavior, etc.	430	3.14
Protect user data, user personal data, data abused activities, unethical experiments with user data, etc.	412	2.79
Personalized content, monetization of user content, personalized messages, etc.	371	2.23
Privacy concerns, digital privacy, privacy-driven access, privacy protection, pricing privacy, privacy score, etc.	332	2.04
Social ads, social engagement, social media profiles, social networks preferences, social networks abuse, etc.	293	1.93
Information management, decision making in management, insights, innovation in management, etc.	257	1.45
Social media trust, trust in the platform, trust in the social media algorithm, user trust to other users, etc.	193	1.08

Source: the authors.

The total grouping of words in nodes was represented by the weighted percentage (WP) that reflected the total weight based on the relevance of the set of keywords in the entire initial database (Hutchison, Johnston, & Breckon, 2010). The analysis was performed using the NVivo Pro-11 textual analysis software with extensions for content filtering and classification.

4. Results

The results of applying the LDA analysis showed a total of 14 main topics in the interview data. Table 12 provides further detail on the identified topics, including also their keyness values and statistical significance (*p*-value).

Furthermore, Table 13 shows the words classified as relevant within the analyzed database. The keywords were grouped based on the nodes analyzed using textual analysis approach. We also report the number of times that the keywords and their synonyms were repeated in the data, as well as their corresponding weight in the entire database.

5. Discussion, implications, and research agenda

In the present study, we identified different challenges related to the implementation and development of DDI strategies that focus on the analysis of user data in social networks and digital markets. Our results suggest that, using DDI, companies personalize their messages based on the needs of their customers. Corresponding algorithms focus on innovation in terms of collecting information from users, allowing companies to find a multitude of data points to predict both user behavior and their actions in digital ecosystems (Sheehan, 2002).

However, the effectiveness of these innovation-focused approach strategies raises privacy concerns (Dutta & Bhat, 2016). Pursuing economic and business objectives (Keller et al., 2017), companies can achieve change in user behavior, or behavioral modification (Zuboff, 2019) based on the application of DDI (Imran-Daud et al., 2016). For instance, psychographic variables and their collection with Big Data techniques allow companies to predict the personality of users. In this respect, our findings are consistent with those reported by Paine et al. (2007) and Qian, F., Ruan, Chen, and Tang (2016).

One of the interviewed informants indicated:

> "*We can say that personality drives user behavior online, and behavior influences user actions in digital markets. These actions generate and mark the personality of the profiles that are then analyzed using DDI*".

In line with Sledgianowski and Kulviwat (2009) and Yu et al. (2019), our results demonstrate that these actions make it possible to understand the factors related to personal data, such as user personality, tastes, habits, and actions in digital environments. Therefore, we can conclude that mining such details from UGD in digital markets allows companies to increase profitability of their content marketing strategies, as they know users better and can personalize content automatically (Brighi, Lucarelli, & Venturelli, 2019; Prince, 2018).

However, as discussed previously, the fact that companies can use DDI techniques to construct psychological profiles of users can lead to unethical experiments that violate the privacy of personal data of users. As noted by one of our interviewees:

> "*We train models that work with machine-learning using common patterns among the users with whom we carry out A/B tests, and on the results of these, we add more information on the profiles used until we achieve the level of accuracy that we consider profitable*".

Although Big Data marketing and content customization make it possible to quantitatively evaluate the effectiveness of advertising and social ads campaigns in digital environments, they also lead to the emergence of trends such as fake news and abusive activities related to advertising (Lies, 2019; Liu & Terzi, 2010).

Therefore, user privacy concern increases as innovation and models developed to get increasingly advanced, and users can perceive these technologies in the personalization of content (see also Palos-Sanchez et al., 2019). Content customization, i.e., segmented targeting based on data-driven models and testing of new market segments, includes actions focused on the analysis of user data applying innovation models and algorithms that study their online actions (He, Bahirat, Knijnenburg, & Menon, 2020). In this respect, one of our interviewees noted:

> "*In the innovation processes, we establish to extract insights that help improve results, we use data we already have on users, but we consider how to request additional information from users, without being intrusive, which will help to improve the accuracy of our models*".

In line with Yu et al. (2019), our results also suggest that the large-scale analysis of user data has led to the massive monetization of users' personal information. Accordingly, there have been concerns about the adverse impact of behavior modifications achieved through abusive privacy practices. In this relation, one of our informants stated:

> "*We have come to question whether the predictive capabilities of our models can influence purchasing behavior and the choices that users make online. We respect user privacy, but the segmentation tools are becoming more robust and intelligent*".

Table 14
Future research questions on user generated-data analysis using data-driven innovation.

User data	Data-driven innovation tools	Future research questions
Intentionally versus non-intentionally generated data	Data-driven models and user data points	• Is it ethical to collect and analyze non-intentionally generated data using data-driven models? • Will such analysis violate user privacy?
Monetization of user content	Data-driven innovations strategies to increase profitability	• What factors influence DDI to increase profitability using data intentionally created by users? • What are the limits of application of DDI in digital markets to obtain the maximum economic return from user-generated content?
Social ads and personalized content	Data-driven innovation actions to personalized social ads	• What is the impact of BDA and DDIs on user behavior when interacting with social ads? • How does the automatic study of the psychographic variables of users using DDI tools affect their purchase decisions in digital environments?
Data abuse activities	Data-driven models to collect and process user information on a large scale	• What framework regulates the limits of predicting user actions in digital markets? • How can large-scale data automation and DDI avoid data abuse activities in automatic or machine-learning models?
Online user behavior	Data-driven innovation based on user online habits	• How can tracking online user behavior influence the decisions that other users make? • Could the development of DDI, focused on understanding the behavior of online users, modify the behavior of online users?
Information management	Decision-making related to the application of strategies based on DDI and artificial intelligences	• How can senior managers of companies that work with user data apply DDI based on artificial intelligences, without violating user privacy? • What factors influence decision making regarding the management, sale, and marketing of user personal data?
Laws on digital privacy	Data-driven innovation models to study online user profiles	• Is it possible to establish a legal framework so that users know the value of their data? • What is the value of user data based on their use of digital markets, social media profiles, and so forth?

Furthermore, several authors argued that social networks and digital environments have a social mission to create online communities to socialize users and strengthen or create social ties among them (e.g., Isaak & Hanna, 2018; Zuboff, 2019). At the same time, other studies argued that it is necessary to strengthen the legislation that regulates the use of targeting tools. Overall, we agree with the Ceron (2015) argument on the need to change the paradigm in social markets.

Although large-scale applications of the tools focused on data innovation can adversely affect users' feelings and privacy, these algorithms ensure the success of companies' communication and marketing strategies (Hajli, 2014). At the same time, there is evidence that digital ecosystems generate addictions in some users (Hawi & Samaha, 2017), and that this addiction can have a negative impact on users' psychological states (Judson et al., 2012). In this context, how would data-driven innovation affect user behavior in the future?

One of the challenges is to create DDI strategies that would prioritize user privacy and interests, rather than companies' profit-driven goals. As specified by one of our interviewees:

> "*What is important is to understand how users must have their own determination to control their data and even to know what the price of the data is. After that, companies will be able to adapt the new ecosystem that protects users and the data they generate on the Internet".*

Therefore, we can conclude that, in the future, data-centric approaches should be able to build marketing models based on ethical design. Corresponding regulation around digital privacy should be developed and introduced. In the context of the rapid expansion of technology and innovation, the current initiatives of the European Commission in the European Union—such as the right to forget, or the new law of General Data Protection Regulation (GDPR) and its impact globally—are clearly not sufficient to fully address all emerging concerns, or any other regulation initiatives worldwide (Goddard, 2017).

In summary, based on the results of the present study and following Bandara, Fernando, and Akter (2019), Saura (2020) and Bandara, Fernando, and Akter (2020) we formulate the following agenda for further research on using DDI strategies to analyze UGD (see Table 14).

5.1. Research propositions to address the challenges and opportunities of UGD analysis using DDI

In order to guide future research in this area and following Hughes et al. (2019) and Duan, Edwards, and Dwivedi (2019), in what follows, we formulate several research propositions based on our results. The proposed propositions are aligned with the categories of DDI models and tools shown in Table 14 that are the results of the literature review and the framework consulted to establish the research theoretical underpinnings.

In addition, these categories of DDI are linked to the objectives of the research to understand how user privacy should be understood from the development of DDI strategies. In the future, these proposals can be used by researchers or practitioners as a starting point for future research and practice in the industry of information management, digital marketing, BDA, and so forth.

Following Li et al. (2020), it is necessary to understand the importance of predictive capacity of new algorithms that work with Artificial Intelligence, as well as to collect and analyze the data of online users. Also, according to Cui and Curry (2005), the more these algorithms are trained, the better are their predictive capabilities, which can lead to decisions focused on economic objectives that anticipate or modify the decisions of users in digital markets. These automations in the collection, analysis, and prediction of online user behavior can lead to privacy violations (Ma, Chen, & Zhang, 2019). Therefore, the following research proposition is set:

Proposition 1. *The ability to collect, analyze, and predict user actions based on the results of the analysis of intentionally and non-intentionally generated data on social networks can violate user's privacy*

The monetization of user actions on the Internet has been one of the digital strategies that has recently evolved in the business environment (Tang, 2016). As argued by Nisar and Yeung (2018), monetization in economic terms of the UGD is key to the profitability of digital business models, as it improves products and services, decision-making, and understanding of the audience (Saura, 2020). Therefore, and following Trabucchi and Buganza (2019), the ability to collect large amounts of

UGD analyzed with DDI models is key for digital business models. Accordingly, the following proposition is established:

Proposition 2. *The greater the predictability and size of the UGD databases, the greater the profitability and monetization of the value of users using DDI models.*

According to Missaglia et al. (2017), the improvement of the study and optimization of social ads and personalized content in digital markets plat a key role in purchase considerations. Holmlund et al. (2020) indicated that the development and application of techniques focused on BDA to explore and influence the customer journey of users is decisive for the success of digital strategies. Accordingly, the study of online user behavior has become a priority for companies that develop digital strategies (Oestreicher-Singer & Zalmanson, 2013). Therefore, the following research proposal is formulated:

Proposition 3. *The use of BDA and DDI for the study of user behavior improves the personalization of social ads / content, increasing the possibilities of positive purchase decisions in digital markets.*

However, the application of DDI to large UGD databases has become a problem for the industry in terms of user perception (Xie & Karan, 2019). As highlighted by Tan, Qin, Kim, and Hsu (2012), one of the challenges is to understand the limits of large-scale automation with the use of DDI, since the prediction and optimization capabilities increase. This effectiveness has led to privacy concerns about the use of the information published by users, as well as the insights, both direct and indirect, which can be extracted from user publications online actions (Huertas & Marine-Roig, 2015). Therefore, the following proposition is formulated

Proposition 4. *Strategies focused on large-scale data automation and DDI must be standardized and examined to avoid abuse that could harm user privacy and data.*

The application of DDI and BDA to the study of online user behavior has been studied from behavioral (Pachidi, Spruit, & Van De Weerd, 2014) and marketing perspectives (Vinerean, Cetina, Dumitrescu, & Tichindelean, 2013; Palos-Sanchez et al., 2019). However, these analytical approaches have allowed tracking users online, allowing thise companies to anticipate user decisions and understand how users behave on the Internet (Steinfeld, 2016; Tene & Polenetsky, 2012). Therefore, and from the point of view of modifying the decisions that users make in digital markets using DDI models, the following proposition is proposed.

Proposition 5. *Tracking online user behavior and using DDI to personalize content and advertising in digital marketplaces may result in the change of decisions that users make in digital environments.*

Information management in this digital age is a key element needed for business success (Dwivedi, Lal, & Williams, 2009). According to Kache and Seuring (2017), in this new connected paradigm, decision-making processes driven by data dashboards is key in marketing, sales, communication, and strategy (Jones, Ball, & Ekmekcioglu, 2008). However, business managers should carefully consider the limits of the use of personal data information in the predictions made to personalized content and increase the benefits. Therefore, the following research proposal is proposed:

Proposition 6. DDI that works with Artificial Intelligence plays an important role in *information management; however, with regard to marketing and sales, the limits of user personal data analysis and predictions should be considered.*

The evolution of DDI in companies and the data generated daily have led to the emergence of a new ecosystem where data are the center of all decisions and strategies implemented in digital markets (Calvano & Polo, 2020). However, according to Morse and Birnhack (2020), the laws on digital privacy have not advanced at a comparable speed. Therefore, since user data and information are used to increase the profits of companies, regulations on the use of user data and information they share on the Internet must be improved (Romansky, 2019). Therefore, the following proposition is formulated:

Proposition 7. *A legal framework to make users aware of the economic value of their data that companies can use to increase their profits should be introduced.*

6. Conclusions

In the present study, we analyzed how, from the perspective of UGD, data-driven models can be used to address the issues of user data privacy.

With regard to our main research question ("What are the challenges of DDI models in digital markets in the context of increasing *user privacy concerns*?), we proposed a detailed research agenda, including the main questions and research propositions that should be addressed in future research regarding using DDI models and strategies with respect to user privacy.

This roadmap for future research is based on the results of our achieving the specific goals of the present investigation. Specifically, we identified definitional perspectives of user privacy in DDI from the UGD theoretical perspective, explored the types of DDI approaches to preserve user privacy in digital markets, reviewed and analyzed what is known about the use of UGD in DDI preserving user privacy, and provided guidelines to track the challenges of DDI with regard to user privacy. Therefore, seven data-driven based topics were found as the main factor to determine next studies in this area of research: intentionally versus non-intentionally generated data, monetization of user content, social ads and personalized content, data abuse activities, online user behavior, information management and laws on digital privacy.

Similarly, DDI tools were found to drive these new challenges: data-driven models and user data points analysis, DDI strategies to increase profitability, data-driven innovation actions to personalized social ads, data-driven models to collect and process user information on a large scale, data-driven innovation based on user habits online, decision-making related to the application of strategies based on DDI, and artificial intelligences DDI models to study online user profiles.

We also reviewed the main uses of the DDI strategies by companies and their link to the privacy of users (in terms of their personality, behavior, and actions on the Internet). Taken together, our results highlight the urgent need to better understand the DDI strategies that could affect user privacy.

6.1. Theoretical contributions

In terms of theoretical implications, the present study provides an adequate framework in relation to the concepts of UGD and DDI for further research on management, processing, and prediction of user behaviors on the Internet based on the data users share in digital markets. Accordingly, future studies can address the questions included in the proposed research agenda.

From the theoretical perspective, researchers should focus on the development of legislation that would regulate the use of targeting tools in digital ecosystems. These initiatives should protect users from abusive privacy practices developed by companies that collect UGC and UGB data from online users.

In addition, future large-scale analyses of user data should follow the best practice guidelines that ensure the appropriate ethical design of both the ways of collecting data and predicting user behavior. In this way, although the economic objectives of companies could be ambitions, companies should ensure that user privacy, the strategies used to influence user online behavior, and predictions about their actions are not violated or abused.

Table A1
Interview questions.

Questions	Codification
What is your use of Data-Driven Innovation (DDI) in your organization?	QD1
What is the role of user-generated data (UGD) in your organization?	QD2
What kind of user-generated data (UGD) do you collect?	QD3
Do you apply DDI-centric models on UGD databases?	QD4
What actions do you take to ensure the privacy of users and their data?	QD5
What use will you give to DDI-based strategies in the future?	QD6
What is the role that DDI and UGD play in the marketing, communication, and data management decisions of the organization?	QD7

6.2. Managerial contributions

From a more practical point of view, managers and heads of communication, marketing, data and development innovation strategies can use the results of the present study as the starting point to develop ethical approaches to the management and processing of user data that would not violate user privacy and appropriately handle user personal information and behavioral data.

In addition, when applying DDI models on these databases, governments, public institutions, and private companies that collect, process, and analyze user data must ensure that user privacy is maintained. With the development and improvement of data science techniques, technology is advancing exponentially; however, no comparable advances are observed in relevant legislation.

Therefore, from the practical and management points of view, it is important that policy makers and managers develop flexible. This is needed to both protect user privacy and to implement DDI strategies that do not infringe user rights.

6.3. Future research and limitations

The limitations of the present study are related to the number of articles identified and reviewed in our systematic literature review, the number of interviewees who participated in the interviews, and the types of analysis used to analyze the data.

In terms of future research objectives, the research propositions described above should be taken into account as starting points to establish new directions and lines of research focused on gaining a better understanding of user behavior with DDI strategies and models.

Authors statement

Conceptualization: J.R.S, D.R.S, D.P.M; Formal analysis: D.P.M; Investigation: J.R.S, D.R.S, D.P.M; Methodology: J.R.S, D.R.S, D.P.M; Resources: J.R.S; Software: J.R.S, D.R.S, D.P.M; Supervision: D.R.S, D.P.M; Validation: J.R.S; Visualization: J.R.S; Roles/Writing - original draft: J.R.S, D.R.S, D.P.M; Writing - review & editing: J.R.S, D.R.S, D.P.M.

Annex 1

Table A1

References

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194. https://doi.org/10.1017/S0963180114000589.

Akter, S., Bandara, R., Hani, U., Wamba, S. F., Foropon, C., & Papadopoulos, T. (2019). Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management, 48*, 85–95. https://doi.org/10.1016/j.ijinfomgt.2019.01.020.

Allen, D. K., & Shoard, M. (2005). Spreading the load: Mobile information and communications technologies and their effect on information overload. *Information Research: An International Electronic Journal, 10*(2), n2.

Amin, A., Shah, B., Khattak, A. M., Moreira, F. J. L., Ali, G., Rocha, Á., et al. (2019). Cross-company customer churn prediction in telecommunication: A comparison of data transformation methods. *International Journal of Information Management, 46*, 304–319. https://doi.org/10.1016/j.ijinfomgt.2018.08.015.

Arya, V., Sethi, D., & Paul, J. (2019). Does digital footprint act as a digital asset? ? Enhancing brand experience through remarketing. *International Journal of Information Management, 49*, 142–156. https://doi.org/10.1016/j.ijinfomgt.2019.03.013.

Areepattamannil, S., & Santos, I. M. (2019). Adolescent students' perceived information and communication technology (ICT) competence and autonomy: Examining links to dispositions toward science in 42 countries. *Computers in Human Behavior, 98*, 50–58. https://doi.org/10.1016/j.chb.2019.04.005.

Bacq, S., Janssen, F., & Noël, C. (2019). What happens next? A qualitative study of founder succession in social enterprises. *Journal of Small Business Management, 57*(3), 820–844. https://doi.org/10.1111/jsbm.12326.

Baird, D. E., & Fisher, M. (2005). Neomillennial user experience design strategies: Utilizing social networking media to support "always on" learning styles. *Journal of Educational Technology Systems, 34*(1), 5–32. https://doi.org/10.2190/6WMW-47L0-M81Q-12G1.

Bandara, R., Fernando, M., & Akter, S. (2019). Privacy concerns in E-commerce: A taxonomy and a future research agenda. *Electronic Markets*, 1–19. https://doi.org/10.1007/s12525-019-00375-6.

Bandara, R., Fernando, M., & Akter, S. (2020a). Managing consumer privacy concerns and defensive behaviours in the digital marketplace. *European Journal of Marketing.* https://doi.org/10.1108/EJM-06-2019-0515. Vol. ahead-of-print No. ahead-of-print.

Bandara, R., Fernando, M., & Akter, S. (2020b). Explicating the privacy paradox: A qualitative inquiry of online shopping consumers. *Journal of Retailing and Consumer Services, 52*, 101947. https://doi.org/10.1016/j.jretconser.2019.101947.

Bandara, R., Fernando, M., & Akter, S. (2020c). Addressing privacy predicaments in the digital marketplace: A power-relations perspective. *International Journal of Consumer Studies, 44*(5), 423–434. https://doi.org/10.1111/ijcs.12576.

Blei, D. M., Ng, A. Y., Jordan, M. I., & Lafferty, J. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993.

Bouncken, R. B., Kraus, S., & Roig-Tierno, N. (2019). Knowledge-and innovation-based business models for future growth: Digitalized business models and portfolio considerations. *Review of Managerial Science*, 1–14.

Boyd, D., Golder, S., & Lotan, G. (2010). Tweet, tweet, retweet: Conversational aspects of retweeting on twitter. In *2010 43rd Hawaii International Conference on System Sciences* (pp. 1–10). https://doi.org/10.1109/HICSS.2010.412.

Brighi, P., Lucarelli, C., & Venturelli, V. (2019). Predictive strength of lending technologies in funding smes. *Journal of Small Business Management, 57*(4), 1350–1377. https://doi.org/10.1111/jsbm.12444.

Büschken, J., & Allenby, G. M. (2016). Sentence-based text analysis for customer reviews. *Marketing Science, 35*(6), 953–975. https://doi.org/10.1287/mksc.2016.0993.

Calvano, E., & Polo, M. (2020). Market power, competition and innovation in digital markets: A survey1. *Information Economics and Policy*, Article 100853.

Ceron, A. (2015). Internet, news, and political trust: The difference between social media and online media outlets. *Journal of Computer-Mediated Communication, 20*(5), 487. https://doi.org/10.1111/jcc4.12129.

Chen, X., Vorvoreanu, M., & Madhavan, K. (2014). Mining social media data for understanding students' learning experiences. *IEEE Transactions on Learning Technologies, 7*(3), 246–259. https://doi.org/10.1109/TLT.2013.2296520.

Cheng, X., Fu, S., & de Vreede, G. J. (2017). Understanding trust influencing factors in social media communication: A qualitative study. *International Journal of Information Management, 37*(2), 25–35. https://doi.org/10.1016/j.ijinfomgt.2016.11.009.

Cheung, M., & She, J. (2016). Evaluating the privacy risk of user-shared images. *ACM Transactions on Multimedia Computing Communications and Applications, 12*(4s), 1–21. https://doi.org/10.1145/2978568.

Cooke-Davies, T. J., & Arzymanow, A. (2003). The maturity of project management in different industries: An investigation into variations between project management models. *International Journal of Project Management, 21*(6), 471–478.

Cui, D., & Curry, D. (2005). Prediction in marketing using the support vector machine. *Marketing Science, 24*(4), 595–615.

de Camargo Fiorini, P., Seles, B. M. R. P., Jabbour, C. J. C., Mariano, E. B., & de Sousa Jabbour, A. B. L. (2018). Management theory and big data literature: From a review to a research agenda. *International Journal of Information Management, 43*, 112–129. https://doi.org/10.1016/j.ijinfomgt.2018.07.005.

Dhillon, G., & Torkzadeh, G. (2006). Value-focused assessment of information system security in organizations. *Information Systems Journal, 16*(3), 293–314.

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management, 48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

Dutta, N., & Bhat, A. (2016). Exploring the effect of store characteristics and interpersonal trust on purchase intention in the context of online social media marketing. *Journal of Internet Commerce, 15*(3), 239–273. https://doi.org/10.1080/15332861.2016.1191053.

Dwivedi, Y. K., Lal, B., & Williams, M. D. (2009). Managing consumer adoption of broadband: Examining drivers and barriers. *Industrial Management & Data Systems*.

Ferreira, J. J., & Teixeira, A. A. (2019). Open innovation and knowledge for fostering business ecosystems. *Journal of Innovation & Knowledge, 4*(4), 253–255. https://doi.org/10.1016/j.jik.2018.10.002.

García-Cabrera, A. M., García-Soto, M. G., & Olivares-Mesa, A. (2019). Entrepreneurs' resources, technology strategy, and new technology-based firms' performance. *Journal of Small Business Management, 57*(4), 1506–1530.

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research, 59*(6), 703–705. https://doi.org/10.2501/IJMR-2017-050.

González, R. J. (2017). Hacking the citizenry?: Personality profiling,'big data'and the election of Donald Trump. *Anthropology Today, 33*(3), 9–12. https://doi.org/10.1111/1467-8322.12348.

Gutierrez, A., O'Leary, S., Rana, N. P., Dwivedi, Y. K., & Calle, T. (2019). Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: Identifying intrusiveness as the critical risk factor. *Computers in Human Behavior, 95*, 295–306. https://doi.org/10.1016/j.chb.2018.09.015.

Hajli, M. N. (2014). A study of the impact of social media on consumers. *International Journal of Market Research, 56*(3), 387–404. https://doi.org/10.2501/IJMR-2014-025.

Hansen, J. M., Saridakis, G., & Benson, V. (2018). Risk, trust, and the interaction of perceived ease of use and behavioral control in predicting consumers' use of social media for transactions. *Computers in Human Behavior, 80*, 197–206. https://doi.org/10.1016/j.chb.2017.11.010.

Hao, H., Zhang, K., Wang, W., & Gao, G. (2017). A tale of two countries: International comparison of online doctor reviews between China and the United States. *International Journal of Medical Informatics, 99*, 37–44. https://doi.org/10.1016/j.ijmedinf.2016.12.007.

Hargittai, E. (2010). Digital na (t) ives? Variation in internet skills and uses among members of the "net generation". *Sociological Inquiry, 80*(1), 92–113. https://doi.org/10.1111/j.1475-682X.2009.00317.x.

Hawi, N. S., & Samaha, M. (2017). The relations among social media addiction, self-esteem, and life satisfaction in university students. *Social Science Computer Review, 35*(5), 576–586. https://doi.org/10.1177/0894439316660340.

He, Y., Bahirat, P., Knijnenburg, B. P., & Menon, A. (2020). A data-driven approach to designing for privacy in household IoT. *ACM Transactions on Interactive Intelligent Systems, 10*(1), 1–47. https://doi.org/10.1145/3241378.

Hilal, A. H., & Alabri, S. S. (2013). Using NVivo for data analysis in qualitative research. *International Interdisciplinary Journal of Education, 2*(2), 181–186. https://doi.org/10.12816/0002914z.

Holmlund, M., Van Vaerenbergh, Y., Ciuchita, R., Ravald, A., Sarantopoulos, P., Ordenes, F. V., et al. (2020). Customer experience management in the age of big data analytics: A strategic framework. *Journal of Business Research*.

Huertas, A., & Marine-Roig, E. (2015). Destination brand communication through the social media: What contents trigger most reactions of users?. *Information and communication technologies in tourism 2015* (pp. 295–308). Cham: Springer.

Hughes, L., Dwivedi, Y. K., Misra, S. K., Rana, N. P., Raghavan, V., & Akella, V. (2019). Blockchain research, practice and policy: Applications, benefits, limitations, emerging research themes and research agenda. *International Journal of Information Management, 49*, 114–129. https://doi.org/10.1016/j.ijinfomgt.2019.02.005.

Hutchison, A. J., Johnston, L. H., & Breckon, J. D. (2010). Using QSR-NVivo to facilitate the development of a grounded theory project: An account of a worked example. *International Journal of Social Research Methodology, 13*(4), 283–302. https://doi.org/10.1080/13645570902996301.

Imran-Daud, M., Sánchez, D., & Viejo, A. (2016). Privacy-driven access control in social networks by means of automatic semantic annotation. *Computer Communications, 76*, 12–25. https://doi.org/10.1016/j.comcom.2016.01.001.

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer, 51*(8), 56–59. https://doi.org/10.1109/MC.2018.3191268.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideologya social identity perspective on polarization. *Public Opinion Quarterly, 76*(3), 405–431. https://doi.org/10.1093/poq/nfs038.

Jiang, B., Liang, J., Sha, Y., Li, R., Liu, W., Ma, H., et al. (2016). Retweeting behavior prediction based on one-class collaborative filtering in social networks. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 977–980). https://doi.org/10.1145/2911451.2914713.

Jimenez-Marquez, J. L., Gonzalez-Carrasco, I., Lopez-Cuadrado, J. L., & Ruiz-Mezcua, B. (2019). Towards a big data framework for analyzing social media content. *International Journal of Information Management, 44*, 1–12. https://doi.org/10.1016/j.ijinfomgt.2018.09.003.

Jones, S., Ball, A., & Ekmekcioglu, Ç. (2008). The data audit framework: A first step in the data management challenge. *International Journal of Digital Curation, 3*(2).

Judson, K. M., Devasagayam, P. R., & Buff, C. L. (2012). Self-perceived brand relevance of and satisfaction with social media. *Marketing Management Journal, 22*(2), 131–144.

Kache, F., & Seuring, S. (2017). Challenges and opportunities of digital information at the intersection of Big Data Analytics and supply chain management. *International Journal of Operations & Production Management*.

Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research–Moving away from the "What" towards the "Why. *International Journal of Information Management, 54*, Article 102205. https://doi.org/10.1016/j.ijinfomgt.2020.102205.

Karegar, F., Pettersson, J. S., & Fischer-Hübner, S. (2020). The dilemma of user engagement in privacy notices. *ACM Transactions on Privacy and Security, 23*(1), 1–38. https://doi.org/10.1145/3372296.

Kaufhold, M. A., Rupp, N., Reuter, C., & Habdank, M. (2020). Mitigating information overload in social media during conflicts and crises: Design and evaluation of a cross-platform alerting system. *Behaviour & Information Technology, 39*(3), 319–342. https://doi.org/10.1080/0144929X.2019.1620334.

Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2017). How to manipulate social media: Analyzing political astroturfing using ground truth data from South Korea. In *Eleventh International AAAI Conference on Web and Social Media*.

Kiss, A. N., Williams, D. W., & Houghton, S. M. (2013). Risk bias and the link between motivation and new venture post-entry international growth. *International Business Review, 22*(6), 1068–1078. https://doi.org/10.1016/j.ibusrev.2013.02.007.

Kraus, S., Breier, M., & Dasí-Rodríguez, S. (2020). The art of crafting a systematic literature review in entrepreneurship research. *International Entrepreneurship and Management Journal*, 1–20.

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed., pp. 221–250). Thousand Oaks, CA, USA: Sage. https://doi.org/10.2307/2288384.

Lee, T. Y., & Bradlow, E. T. (2011). Automated marketing research using online customer reviews. *Journal of Marketing Research, 48*(5), 881–894. https://doi.org/10.2139/ssrn.1726055.

Li, S., Song, X., Lu, H., Zeng, L., Shi, M., & Liu, F. (2020). Friend recommendation for cross marketing in online brand community based on intelligent attention allocation link prediction algorithm. *Expert Systems With Applications, 139*, Article 112839.

Lies, J. (2019). Marketing intelligence and big data: Digital marketing techniques on their way to becoming social engineering techniques in marketing. *International Journal of Interactive Multimedia and Artificial Intelligence, 5*(5). https://doi.org/10.9781/ijimai.2019.05.002.

Liu, T., & Lai, Z. (2020). From non-player characters to othered participants: Chinese women's gaming experience in the 'free'digital market. *Information, Communication and Society*, 1–19. https://doi.org/10.1080/1369118X.2020.1791217.

Liu, K., & Terzi, E. (2010). A framework for computing the privacy scores of users in online social networks. *ACM Transactions on Knowledge Discovery From Data, 5*(1), 1–30. https://doi.org/10.1145/1870096.1870102.

Lozano, M. G., Schreiber, J., & Brynielsson, J. (2017). Tracking geographical locations using a geo-aware topic model for analyzing social media data. *Decision Support Systems, 99*, 18–29. https://doi.org/10.1016/j.dss.2017.05.006.

Ma, X., Chen, X., & Zhang, X. (2019). Non-interactive privacy-preserving neural network prediction. *Information Sciences, 481*, 507–519.

MacDougall, C., & Fudge, E. (2001). Planning and recruiting the sample for focus groups and in-depth interviews. *Qualitative Health Research, 11*(1), 117–126. https://doi.org/10.1177/104973201129118975.

Malgieri, G., & Custers, B. (2018). Pricing privacy – The right to know the value of your personal data. *Computer Law & Security Report, 34*(2), 289–303. https://doi.org/10.1016/j.clsr.2017.08.006.

Martinelli Watanuki, H., & de Oliveira Moraes, R. (2019). Exploring the influence of social media information on interpersonal trust in new virtual work partners. *Informatics, 6*(September (3)). https://doi.org/10.3390/informatics6030033.

Martinez, M. G., & Walton, B. (2014). The wisdom of crowds: The potential of online communities as a tool for data analysis. *Technovation, 34*(4), 203–214. https://doi.org/10.1016/j.technovation.2014.01.011.

McKnight, D. H., Kacmar, C. J., & Choudhury, V. (2004). Dispositional trust and distrust distinctions in predicting high-and low-risk internet expert advice site perceptions. *E-Service, 3*(2), 35–58. https://doi.org/10.2979/ESJ.2004.3.2.35.

Minhas, S., & Hussain, A. (2014). Linguistic correlates of deception in financial text a corpus linguistics based approach. *Psychology Review, 19*, 307–342.

Missaglia, A. L., Oppo, A., Mauri, M., Ghiringhelli, B., Ciceri, A., & Russo, V. (2017). The impact of emotions on recall: An empirical study on social ads. *Journal of Consumer Behaviour, 16*(5), 424–433.

Morse, T., & Birnhack, M. (2020). The posthumous privacy paradox: Privacy preferences and behavior regarding digital remains. *New Media & Society*, Article 1461444820974955.

Ndumu, A. (2019). Linkages between information overload and acculturative stress: The case of Black diasporic immigrants in the US. *Journal of Librarianship and Information Science.* doi: 0961000619857115.

Netzer, Y., Tenenboim-Weinblatt, K., & Shifman, L. (2014). The construction of participation in news websites: A five–Dimensional model. *Journalism Studies, 15*(5), 619–631. https://doi.org/10.1080/1461670X.2014.895527.

Nisar, T. M., & Yeung, M. (2018). Attribution modeling in digital advertising: An empirical investigation of the impact of digital sales channels. *Journal of Advertising Research, 58*(4), 399–413.

Öberg, C., & Alexander, A. T. (2019). The openness of open innovation in ecosystems–Integrating innovation and management literature on knowledge linkages. *Journal of Innovation & Knowledge, 4*(4), 211–218. https://doi.org/10.1016/j.jik.2017.10.005.

Oestreicher-Singer, G., & Zalmanson, L. (2013). Content or community? A digital business strategy for content providers in the social age. *MIS Quarterly*, 591–616.

Omilion-Hodges, L. M., & Rodriguez, M. (2014). The effect of system generated cues on users perceptions of organizational trustworthiness, competence and goodwill. *Journal of Media and Communication Studies, 6*(10), 161–164. https://doi.org/10.5897/JMCS2014.0405.

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research*, (2:1), 1–28. https://doi.org/10.1287/isre.2.1.1.

Pachidi, S., Spruit, M., & Van De Weerd, I. (2014). Understanding users' behavior with software operation data mining. *Computers in Human Behavior, 30*, 583–594.

Paine, C., Reips, U. D., Stieger, S., Joinson, A., & Buchanan, T. (2007). Internet users' perceptions of 'privacy concerns' and 'privacy actions'. *International Journal of Human-computer Studies, 65*(6), 526–536. https://doi.org/10.1016/j.ijhcs.2006.12.001.

Palos-Sanchez, P., Saura, J. R., & Martin-Velicia, F. (2019). A study of the effects of Programmatic Advertising on users' Concerns about Privacy overtime. *Journal of Business Research, 96*(2019), 61–72. https://doi.org/10.1016/j.jbusres.2018.10.059.
Panahi, S., Watson, J., & Partridge, H. (2016). Fostering interpersonal trust on social media: Physicians' perspectives and experiences. *Postgraduate Medical Journal, 92* (1084), 70–73. https://doi.org/10.1136/postgradmedj-2015-133270.
Pangrazio, L., & Selwyn, N. (2019). Personal data literacies': A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society, 21*(2), 419–437. https://doi.org/10.1177/1461444818799523.
Peng, G. (2010). Critical mass, diffusion channels, and digital divide. *Journal of Computer Information Systems, 50*(3), 63–71. https://doi.org/10.1080/08874417.2010.11645408.
Prasarnphanich, P., & Wagner, C. (2009). Explaining the sustainability of digital ecosystems based on the wiki model through critical-mass theory. *IEEE Transactions on Industrial Electronics, 58*(6), 2065–2072. https://doi.org/10.1109/TIE.2009.2027248.
Pratesi, F., Gabrielli, L., Cintia, P., Monreale, A., & Giannotti, F. (2020). PRIMULE: Privacy risk mitigation for user profiles. *Data & Knowledge Engineering, 125*, Article 101786. https://doi.org/10.1016/j.datak.2019.101786.
Prince, C. (2018). Do consumers want to control their personal data? Empirical evidence. *International Journal of Human-computer Studies, 110*, 21–32. https://doi.org/10.1016/j.ijhcs.2017.10.003.
Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.
Qi, L., Zhang, X., Li, S., Wan, S., Wen, Y., & Gong, W. (2020). Spatial-temporal data-driven service recommendation with privacy-preservation. *Information Sciences, 515*, 91–102. https://doi.org/10.1016/j.ins.2019.11.021.
Qian, J., Qiu, F., Wu, F., Ruan, N., Chen, G., & Tang, S. (2016). Privacy-preserving selective aggregation of online user behavior data. *IEEE Transactions on Computers*. https://doi.org/10.1109/tc.2016.2595562, 1–1.
Ramirez-Andreotta, M. D., Brody, J. G., Lothrop, N., Loh, M., Beamer, P. I., & Brown, P. (2016). Improving environmental health literacy and justice through environmental exposure results communication. *International Journal of Environmental Research and Public Health, 13*(7), 690. https://doi.org/10.3390/ijerph13070690.
Rayson, P., & Garside, R. (2000). Comparing corpora using frequency profiling. In *The Workshop on Comparing Corpora* (pp. 1–6). https://doi.org/10.3115/1117729.1117730.
Reyes-Menendez, A., Saura, J. R., & Stephen, B. T. (2020). Exploring key indicators of social identity in the #MeToo era: Using discourse analysis in UGC. *International Journal of Information Management, 54*, Article 102129. https://doi.org/10.1016/j.ijinfomgt.2020.102129.
Ricciardi, F., Zardini, A., & Rossignoli, C. (2018). Organizational integration of the IT function: A key enabler of firm capabilities and performance. *Journal of Innovation & Knowledge, 3*(3), 93–107. https://doi.org/10.1016/j.jik.2017.02.003.
Roberts, L. D. (2015). Ethical issues in conducting qualitative research in online communities. *Qualitative Research in Psychology, 12*(3), 314–325. https://doi.org/10.1080/14780887.2015.1008909.
Romansky, R. (2019). A survey of informatization and privacy in the digital age and basic principles of the new regulation. *International Journal on Information Technologies and Security, 11*(1), 95–106.
Rosa, H., Carvalho, J. P., Astudillo, R., & Batista, F. (2015). Detecting user influence in twitter: Pagerank vs. Katz, a case study. In *Proceedings of the Seventh European Symposium on Computational Intelligence and Mathematics* (pp. 7–10).
Sarkis, J., Zhu, Q., & Lai, K. H. (2011). An organizational theoretic review of green supply chain management literature. *International Journal of Production Economics, 130*(1), 1–15. https://doi.org/10.1016/j.ijpe.2010.11.010.
Saura, J. R. (2020). Using data sciences in digital marketing: Framework, methods, and performance metrics. *Journal of Innovation & Knowledge, 1*(2020). https://doi.org/10.1016/j.jik.2020.08.001.
Saxena, D., & Lamest, M. (2018). Information overload and coping strategies in the big data context: Evidence from the hospitality sector. *Journal of Information Science, 44* (3), 287–297. https://doi.org/10.1177/0165551517693712.
Schoen, H., Gayo-Avello, D., Metaxas, P. T., Mustafaraj, E., Strohmaier, M., & Gloor, P. (2013). The power of prediction with social media. *Internet Research*. https://doi.org/10.1108/IntR-06-2013-0115.
Sembada, A. Y., & Koay, K. Y. (2019). How perceived behavioral control affects trust to purchase in social media stores. *Journal of Business Research, 2019*. https://doi.org/10.1016/j.jbusres.2019.09.028.
Sheehan, K. B. (2002). Toward a typology of Internet users and online privacy concerns. *The Information Society, 18*(1), 21–32. https://doi.org/10.1080/01972240252818207.
Sledgianowski, D., & Kulviwat, S. (2009). Using social network sites: The effects of playfulness, critical mass and trust in a hedonic context. *Journal of Computer Information Systems, 49*(4), 74–83.
Spence, P. R., Lachlan, K. A., Westerman, D., & Spates, S. A. (2013). Where the gates matter less: Ethnicity and perceived source credibility in social media health messages. *Howard Journal of Communications, 24*(1), 1–16. https://doi.org/10.1080/10646175.2013.748593.
Steinfeld, N. (2016). "I agree to the terms and conditions": (how) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior, 55*, 992–1000.
Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. *International Journal of Information Management, 39*, 156–168. https://doi.org/10.1016/j.ijinfomgt.2017.12.002.
Su, C., & Contractor, N. (2011). A multidimensional network approach to studying team members' information seeking from human and digital knowledge sources in consulting firms. *Journal of the American Society for Information Science and Technology, 62*(7), 1257–1275. https://doi.org/10.1002/asi.21526.
Szymczak, H., Kücükbalaban, P., Lemanski, S., Knuth, D., & Schmidt, S. (2016). Trusting Facebook in crisis situations: The role of general use and general trust toward Facebook. *Cyberpsychology, Behavior and Social Networking, 19*(1), 23–27. https://doi.org/10.1089/cyber.2015.0450.
Tahir, M., Habaebi, M. H., Dabbagh, M., Mughees, A., Ahad, A., & Ahmed, K. I. (2020). A review on application of blockchain in 5G and beyond networks: Taxonomy, field-trials, challenges and opportunities. *IEEE Access : Practical Innovations, Open Solutions, 8*, 115876–115904. https://doi.org/10.1109/access.2020.3003020.
Tan, X., Qin, L., Kim, Y., & Hsu, J. (2012). Impact of privacy concern in social networking web sites. *Internet Research*.
Tang, A. K. (2016). Mobile app monetization: App business models in the digital era. *International Journal of Innovation Management and Technology, 7*(5), 224.
Tene, O., & Polenetsky, J. (2012). To track or do not track: Advancing transparency and individual control in online behavioral advertising. *Minn. JL Sci. & Tech., 13*, 281.
Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., et al. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports, 7*(1), 1–11. https://doi.org/10.1038/s41598-017-18007-4.
Törnberg, A., & Törnberg, P. (2016). Muslims in social media discourse: Combining topic modeling and critical discourse analysis. *Discourse Context & Media, 13*, 132–142. https://doi.org/10.1016/j.dcm.2016.04.003.
Trabucchi, D., & Buganza, T. (2019). Data-driven innovation: Switching the perspective on Big Data. *European Journal of Innovation Management*.
Tsvere, M., Swamy, S., & Nyaruwata, T. L. (2013). Perceived competence of Zimbabwean academics in the use of information technology in university academic business. *International Journal of Science and Research, 2*(8), 284–289.
Utz, S., & Krämer, N. C. (2009). The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology Journal of Psychosocial Research on Cyberspace, 3*(2).
Vanhala, M., Lu, C., Peltonen, J., Sundqvist, S., Nummenmaa, J., & Järvelin, K. (2020). The usage of large data sets in online consumer behaviour: A bibliometric and computational text-mining–Driven analysis of previous research. *Journal of Business Research, 106*, 46–59. https://doi.org/10.1016/j.jbusres.2019.09.009.
Vinerean, S., Cetina, I., Dumitrescu, L., & Tichindelean, M. (2013). The effects of social media marketing on online consumer behavior. *International Journal of Business and Management, 8*(14), 66.
Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication, 19*(2), 171–183. https://doi.org/10.1111/jcc4.12041.
Wohn, D. Y., Lee, Y. H., Sung, J., & Bjornrud, T. (2010). Building common ground and reciprocity through social network games. *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, 4423–4428. https://doi.org/10.1145/1753846.1754164.
Wu, H., Yue, K., Pei, Y., Li, B., Zhao, Y., & Dong, F. (2016). Collaborative topic regression with social trust ensemble for recommendation in social media systems. *Knowledge-based Systems, 97*, 111–122. https://doi.org/10.1016/j.knosys.2016.01.011.
Xie, W., & Karan, K. (2019). Consumers' privacy concern and privacy protection on social network sites in the era of big data: Empirical evidence from college students. *Journal of Interactive Advertising, 19*(3), 187–201.
Yang, P., Xiong, N., & Ren, J. (2020). Data security and privacy protection for cloud storage: A survey. *IEEE Access : Practical Innovations, Open Solutions, 8*, 131723–131740. https://doi.org/10.1109/access.2020.3009876.
Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior, 27*(2), 634–639. https://doi.org/10.1016/j.chb.2010.04.014.
Yeon Cho, J., Ko, D., & Gyou Lee, B. (2018). Strategic approach to privacy Calculus of wearable device user regarding information disclosure and continuance intention. *Transactions on Internet and Information Systems, 12*(7). https://doi.org/10.3837/tiis.2018.07.020.
Yu, C., Zhang, Z., Lin, C., & Wu, Y. J. (2019). Can data-driven precision marketing promote user ad clicks? Evidence from advertising in WeChat moments. *Industrial Marketing Management*. https://doi.org/10.1016/j.indmarman.2019.05.001.
Zhong, W., Yin, X., Zhang, X., Li, S., Dou, W., Wang, R., et al. (2020). Multi-dimensional quality-driven service recommendation with privacy-preservation in mobile edge environment. *Computer Communications, 157*, 116–123. https://doi.org/10.1016/j.comcom.2020.04.018.
Zhou, L., Wu, D., Wei, X., & Dong, Z. (2019). Seeing isn't believing: QoE evaluation for privacy-aware users. *IEEE Journal on Selected Areas in Communications, 37*(7), 1656–1665. https://doi.org/10.1109/jsac.2019.2916452.
Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75–89. https://doi.org/10.1057/jit.2015.5.
Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. *Profile books*.