Research Article

Stewardship of personal data on social networking sites

Spyros Angelopoulos[a,*], Michael Brown[b], Derek McAuley[c], Yasmin Merali[d], Richard Mortier[e], Dominic Price[c]

[a] *Tilburg School of Economics and Management, Tilburg University, Tilburg, Netherlands*
[b] *Fowlam Ltd, Nottingham, UK*
[c] *Horizon Digital Economy Research Institute, University of Nottingham, Nottingham, UK*
[d] *Hull University Business School, University of Hull, Hull, UK*
[e] *Department of Computer Science and Technology, University of Cambridge, Cambridge, UK*

A B S T R A C T

The extant Information Management literature highlights the asymmetric distribution of power between users and online platforms, while the issues related to the stewardship of personal data on such platforms remain problematic and largely unresolved. To address that lacuna, we propose a conceptual design that can help to overcome many of the challenges related to storage, analysis, and integrity associated with the stewardship of personal data on online platforms. We adopt a *systemic* perspective and propose a shift from the current user-platform relationship to one in which users control the level of access to their data, organisations are relieved from the burden of maintaining personal data, and the data are not decoupled from information about their provenance and context of origin. We apply our conceptual design to the context of social networking sites, where we specifically address issues related to privacy, and identity and pave the path to a broader set of possible applications. We discuss the significance and timeliness of our proposed conceptual design for the stewardship of personal data, and the importance of our findings for future research, as well as for the design of online platforms.

1. Introduction

The escalation in quantity and granularity of personal data can reveal the behaviours, preferences, and personality traits of users on online platforms (Georgiadou, Angelopoulos, & Drake, 2019; Ioannou, Tussyadiah, & Lu, 2020). The ability to manage and leverage personal data is a critical determinant of competitive advantage (Liu, Soroka, Han, Jian, & Tang, 2019), and this imperative has spawned business models associated with data trading (Perera et al., 2017), and the possibility for payments in exchange for personal data (Lu, Ou, & Angelopoulos, 2018). Such opportunities entail addressing integrity, and security issues and bring to the fore issues related to the breach of privacy, the transparency of purpose for which personal data is collected, and the use to which it is put (Ghasemaghaei, 2020).

The Information Management (IM) literature to date has focused on data management problems (e.g. de Camargo Fiorini, Seles, Jabbour, Mariano, & de Sousa Jabbour, 2018; Ghasemaghaei, 2019, 2020), highlighting the asymmetric distribution of power between users and online platforms that collect and exploit personal data (e.g. Ioannou et al., 2020; Spanaki, Gürgüç, Mulligan, & Lupu, 2019), while the

problems related to the stewardship of personal data on such platforms remain timely and largely unresolved (Gregory, Henfridsson, Kaganer, & Kyriakou, 2020). We address that lacuna, by articulating a conceptual design (Nunamaker, Chen, & Purdin, 1991) to overcome the challenges of stewardship of personal data, and to enable the development of privacy-focused platforms (e.g. Ioannou et al., 2020).

We frame our work along the lines of established IM studies (e.g. Conboy, 2009), and follow a design science approach (Gregor & Hevner, 2013; Grønli, Ghinea, & Bygstad, 2013; Hevner, March, Park, & Ram, 2004; Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007). In doing so, we first identify the shortcomings of the stewardship of personal data on online platforms, based on which we then define the objectives for a solution, and propose a conceptual design incorporating them. Specifically, we use the principles of distributed computing (e.g. Bayramusta & Nasir, 2016; Senyo, Addae, & Boateng, 2018) and separation of concerns (SoC) (Hürsch & Lopes, 1995) to address issues related to data storage, analysis, and integrity, whilst enabling online platform users to exercise more control over their data. We adopt a *systemic* perspective and propose a shift from the current user-platform relationship to one in which: i) users control the level of access to their

---

* Corresponding author.
E-mail address: s.angelopoulos@uvt.nl (S. Angelopoulos).

data, ii) organizations are relieved of the burden of maintaining data, and iii) data are not decoupled from information about their provenance and context of origin. Following these, we demonstrate the utility of our conceptual design through an application on social networking sites (SNS), for which we conduct a survey to identify strategies deployed by users to manage their privacy on the various SNS (e.g. Huberman et al., 2005). Specifically, considering recent cases of alleged misuse of SNS personal data, we address issues related to privacy and identity (e.g. Krasnova, Veltri, Eling, & Buxmann, 2017; Shibchurn & Yan, 2015) and complement our conceptual design with two proposed processors for privacy and identity management, paving the path to a broader set of applications that can provide users of online platforms with greater granularity of control over their data.

This study makes an artefactual contribution (Ågerfalk & Karlsson, 2020) to the extant IM literature focusing on privacy issues (Ghasemaghaei, 2019, 2020) and systemic problems (Abbasi, Sarker, & Chiang, 2016; de Camargo Fiorini et al., 2018), by motivating and implementing a conceptual design in response to the challenges related to the stewardship of personal data on online platforms, without compromising users' privacy (Agarwal & Dhar, 2014; Goes, 2014). We also contribute to the IM literature on privacy-focused SNS, by applying our conceptual design to this context to address issues related to privacy and identity (Krasnova et al., 2017), and to pave the path to a broader set of possible applications.

Our paper is organized as follows: in the next section, we delineate the features of personal data and the issues of leveraging their potential, revealing the need to develop an alternative to the current approach of personal data collection, storage, and analysis. We then propose a conceptual design to meet that need, and illustrate its merits by applying it in SNS. In the penultimate section, we present the limitations of the proposed design, as well as its implications for organizations, users, and the society at large concerning a change in public perceptions. We conclude the paper by reflecting on opportunities for further research.

2. Background and problem statement

Following the approach of Peffers et al. (2007), we initially pinpoint the shortcomings of the stewardship of personal data on online platforms. Extending the harvesting of personal data across various aspects of peoples' daily lives can enable organizations to articulate individuals' behaviours, preferences, and identities (Agarwal & Dhar, 2014; Goes, 2014). The collection of individuals' intimate information on online platforms presents a series of significant ethical challenges, the more wicked of which are concerned with:

- The provenance of data and the need to address the relevance of original contextual information when combining data from diverse sources (Montealegre, Hovorka, & Germonprez, 2014),
- The context-dependent immediacy and relevance of the required information, and the contingent nature and complexity of the socially situated *systems-in-use* (Lycett, 2013),
- The legal and ethical issues, and the behavioural inferences related to the exploitation of personal data in ways that the users are not aware of (Abbasi et al., 2016).

Whilst the value proposition for using personal data may be attractive to organizations (2020, Ghasemaghaei, 2019; Liu et al., 2019; Zuo, Angelopoulos, Ou, Liu, & Liang, 2020) and governments (Georgiadou et al., 2019), their use can spawn significant ethical concerns (Janse, Ou, Angelopoulos, Davison, & Jia, 2017; Perera et al., 2017). Novel models for incentivizing and engaging users, challenge conventional positions on issues such as privacy protection, and make use of gamification (Lowry, Gaskin, & Moody, 2015) and monetary or other incentives (Lu et al., 2018) to persuade users to hand over personal data. The terms of service (ToS) force individuals who want to

register on online platforms to agree to their data being used by those platforms and potentially by others contracting with them. Even when reasonable ToS are offered, there is always a possibility that an online platform could merge with another that is less scrupulous with the use of personal data. In terms of the premises of the newly emerged field of Human-Data Interactions (e.g. Mortier et al., 2016; Perera et al., 2017), such models lack legibility, agency, and negotiability, since the users:

- Cannot discern the multiple sources their data are collected from, the analysis that is performed on their data, and the consequences of that analysis,
- Lack awareness regarding the ways that exist to affect data collection and analysis, and
- Are trapped by the imposed features, facing binary accept/reject ToS mechanisms, the difficulty of refining decisions, and of correcting data or inferences held about them.

Concurrently, data regulations around the world (with prominent ones being Europe's GDPR, USA's HIPAA, NIST 800-171, GLBA, FISMA, Canada's PIPEDA, and Australia's APA), shifted the discussion and presented organizations that collect personal data with a new reality—one in which holding personal data can become a liability. Consequently, many organizations that collect personal data have become sensitive to the issue, especially those associated with SNS, which have recently suffered from several personal-data-related liabilities.

The difficulty in quantifying the risk of handling and abusing personal data raises significant privacy concerns (Abbasi et al., 2016). The work of Kramer, Guillory, and Hancock (2014), for instance, highlights a more insidious aspect of the acquisition and deployment of personal data, demonstrating how such data can be used to manipulate emotions without the consent of users. More recently, the alleged misuse of Facebook users' data by Cambridge Analytica illustrates the problem of transitivity of personal data: users entrusted their data to Facebook for a specific set of reasons and benefits, but, without their consent or knowledge, their data were passed on to 3rd parties, who allegedly committed the transgression of exploiting these data for analysis, psychological profiling, targeting change of opinions, and tampering with democratic processes.

These cases bring to the fore the burden that responsible stewardship of personal data imposes on organizations, highlighting the ethical, psychological, legal, economic, and political consequences that follow when the stewardship of personal data is compromised. They also accentuate the need for novel solutions that safeguard the privacy of users on online platforms without stifling the ethical deployment of personal data in business, and society (Georgiadou et al., 2019; Ghasemaghaei, 2019, 2020; Janse et al., 2017; Liu et al., 2019).

These considerations have led to greater public scrutiny of questionable practices regarding the sharing of personal data with 3rd parties and have prompted popular SNS to make promises regarding their safeguarding of personal data. Whilst such promises have failed to materialize, they have also failed to address the fact that personal data remain consolidated and stored on the organization side, creating in this way a 'honey-pot' for malicious entities. As a result, the continuous leak of personal data by such entities remains a serious problem.

Concurrently, SNS users with overlapping networks extending in both the online and offline spaces resort to alternative solutions to self-manage their privacy by using multiple identities (Angelopoulos & Merali, 2015, 2017). In projecting their identities through their SNS interactivity, users demonstrate a small number of identities at a time, based on their situational context, making these exposures context-dependent. Through the self-management of their multiple identities, the users can control how others perceive them within a certain setting (Ellison, Heino, & Gibbs, 2006; Schwämmlein & Wodzicki, 2012). SNS users, thus, may turn into the distinctive affordances and social norms of the various SNS to craft contextually specific presentations of self, and to distribute specific content on specific SNS, possibly influenced

by the affordances as well as the ToS of the SNS (Zhang & Leung, 2015). Such practices, however, impede the development of consolidated profiles, and their attendant risks (Krasnova et al., 2017; Shibchurn & Yan, 2015). To address the aforementioned issues, we frame our study along the lines of established IM research (e.g. Conboy, 2009), and follow a design science approach (Gregor & Hevner, 2013; Grønli et al., 2013; Hevner et al., 2004; Peffers et al., 2007).

3. Methodology

Design science seeks to develop artifacts for solving problems while differing from theory-building and theory-testing, which seek observation-based explanations (Holmström, Ketokivi, & Hameri, 2009). In doing so, design science studies make artefactual contributions to the IM literature (Ågerfalk & Karlsson, 2020), and as such, the approach has attracted the interest of many IM scholars (Gregor & Hevner, 2013; Grønli et al., 2013; Hevner et al., 2004; Peffers et al., 2007). We follow Holmström et al. (2009) and adopt a definition that focuses on exploration by design, in which design science explores novel alternatives to solve a problem, elaborates on this exploration, and ultimately improves problem-solving. Typically, design science studies: i) identify the problem, ii) define the objectives for a solution, iii) propose a design for the solution, iv) demonstrate the design, v) evaluate it, and ultimately vi) communicate the solution (Peffers et al., 2007). Whilst the evaluation phase is crucial in design science (Hevner et al., 2004; Peffers et al., 2007), it is not uncommon to be omitted when the research context does not allow such an opportunity or to be limited by the study conditions (e.g. Spanaki, Karafili, Sivarajah, Despoudi, & Irani, 2020).

Aligning our work to the design science approach (Peffers et al., 2007), we first i) identify the problem, and ii) define the objectives based on the considerations described in our problem statement. We then iii) propose a conceptual design in response to the challenges related to the stewardship of personal data on online platforms, without compromising users' privacy. Following this, iv) we demonstrate the utility of our proposed conceptual design through an application on SNS to address issues related to privacy and identity. To understand the privacy-related needs of SNS users, as well as the strategies they employ to manage their privacy on the various SNS, we conduct a survey, and based on the results, we enhance our conceptual design by proposing further enhancements. Our study is limited by the absence of an evaluation phase, as this would require collaboration with existing SNS in implementing our design, along with the possible business models that could complement it, which is beyond our scope.

4. Conceptual design

The considerations that we described in the problem statement, define the objectives for the development of our conceptual design, which embodies a paradigm shift in how personal data are accessed and used by online platforms. In our proposed conceptual design, personal data reside with the users, allowing them greater control over access to the data (Fig. 1).

An organization wanting to conduct analytics over a particular set of attributes, can locate those users that allow access to the required set, negotiate access to the specific data needed to conduct analytics directly on the privately-held data, and receive back only the results of the required analysis, while the data remain on the user-side throughout this chain of events (Fig. 2). In contrast to traditional approaches based on the selection of representative data, our approach allows organizations to *target* users, and so can be used for more accurate service personalization based on the preferences of users, or for analytics on specific groups of users.

The ownership and possession of data reside with the originators, along with the discretion to choose whether to allow access to specific services. This entails a shift in power from the online platforms that exploit personal data to the originators of the data. Our approach retains the originator's context and information and eliminates the need for centralized management and storage by providing distributed, and granular personal data store maintained by the users. Moreover, the sample selected in traditional statistical analysis is predicated on a frame derived from ontological assumptions about the structure of the world, while currently for personal data analytics the selection of a representative sample is predicated on the intrinsic structure of the data, in an algorithmically determined way (Chang, Kauffman, & Kwon, 2014; Kitchin, 2014). Furthermore, in the traditional approach, the wealth that personal data held is rarely used by organizations, but organizations still bear the burden of maintaining costly infrastructures to retain the data. Our proposed conceptual design eliminates such a need, and provides organizations with a way to directly target specific users or groups of users.

Our design is predicated on the principle of SoC, providing a connecting layer over existing services, to enable a common interface for flexible and scalable access to personal data across a heterogeneous user base. SoC rests on the idea that system elements should have exclusivity of purpose, and there should not be elements sharing responsibilities. It, thus, focuses on delineating system elements to make the system manageable, resilient, and robust. This entails establishing constraints so that the system comprises of elements with non-repeating responsibilities, reducing the overhead for managing the complexity of the system.

Based on the proposed conceptual design, we implement an application architecture which—following the principles of SoC—has three layers: shim, storage, and processing. To account for interoperability, the application functions similarly to an email client, aggregating SNS activity by using secure communication protocols. The application architecture is shown in Fig. 3, where the various SNS are represented as 'Data Source', from which the application updates the 'Shim Layer', which provides compatibility with application programming interfaces (APIs). The data are then transferred to the 'Storage Layer', which resides on the user-side as predicated by our proposed conceptual design (Fig. 1). The proposed architecture is enhanced with a series of 'processors', which are responsible for transforming data with small, agent-like applications before, ultimately, the user receives the information through the client.

Shim layer. Following the principles of SoC in our architecture, we incorporate a shim layer that prevents duplication and provides extensibility, lowering the deployment and maintenance costs. Such a layer ensures compatibility with newer versions of applications and resolves issues during the development of applications for new online platforms. The shim layer provides a connection with the existing SNS
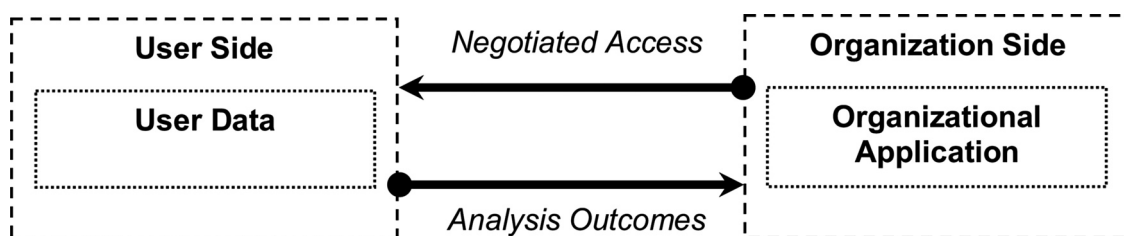

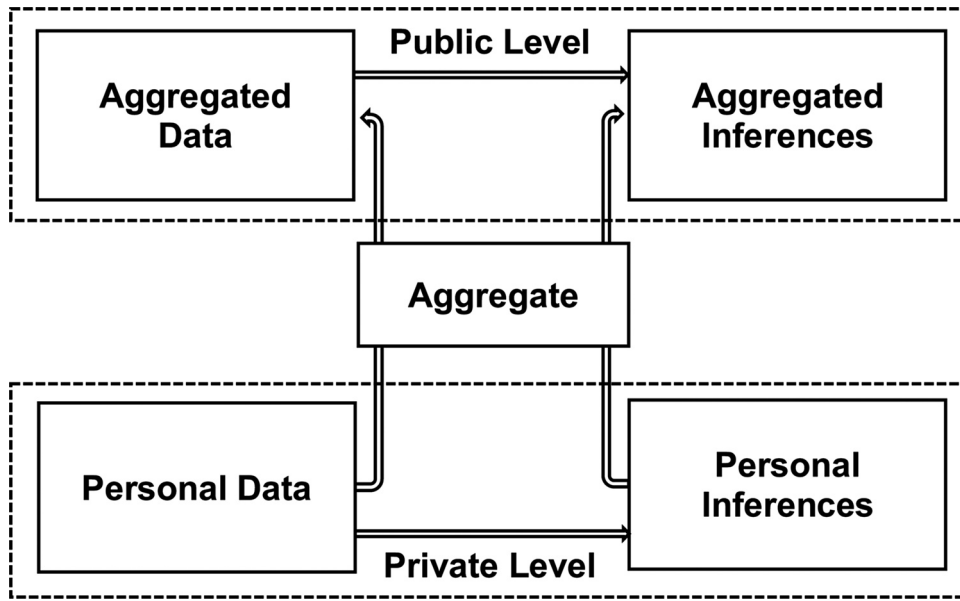**Fig. 1.** Data storage and access in the proposed conceptual design.

Fig. 2. Data analytics in the proposed conceptual design.

to aggregate data into a unified format, and to distribute their transmission. Furthermore, this layer collects the ToS of each SNS, which are then used by the processing layers. Such a layer in our architecture, thus, communicates with the various SNS, and converts the SNS data to a uniform format, making them available to the other layers.

Storage layer. Data storage and redundancy represent a significant element of the proposed conceptual design, since all personal data are stored and replicated on the user-side. The sharing of information through data replication ensures consistency amongst redundant resources and improves reliability, fault-tolerance, and data accessibility. The data aggregated by the shim layer are stored in the storage layer, which is a distributed repository, acting as an intermediary for the user client, releasing the user from the need to communicate directly with the SNS. Such a design enables real-time operation, by receiving the data from the first layer and providing them to the *processors*, while storing them for later use, based on the user preferences. Data such as ToS and past communication, therefore, can be stored and updated only when it is needed, for later use by the processing layer. The storage layer, therefore, enhances resilience by providing transparency, redundancy, and efficacy to the system.

Processing layer. The processing layer provides an input and output interface and is responsible for transforming the data flowing through it with a series of *processors* before users receive the information through a client. This layer is pluggable with published APIs, which can enable 3rd parties to experiment with it and extend the architecture with additional *processors*. Moreover, the processing layer enables control over the presentation of content to the users.

4.1. Processor design: a design science approach to address user requirements

As predicated by the approach of Peffers et al. (2007), we demonstrate the utility of our conceptual design through an application on SNS. To better understand the privacy-related needs of SNS users (Wang, Yan, Lin, & Cui, 2017; Dinev & Hart, 2006; Guo, Lu, Kuang, & Wang, 2020; Malhotra, Kim, & Agarwal, 2004; Sullivan & Koh, 2019; Wang, Duong, & Chen, 2016), as well as the strategies they employ to manage their privacy on the various SNS (e.g. Huberman et al., 2005), we conducted an online survey, exploring i) how users behave on the various SNS, ii) how they self-manage their multiple context-dependent identities, iii) how much and what they reveal about their lives on those SNS, and iv) how they manage their online social networks overall. Following Peffers et al. (2007), and considering cultural influences on users' perception and preferences when initiating a system design, we constructed a questionnaire based on the relevant literature (Dholakia, Bagozzi, & Pearo, 2004; Karl & Peluchette, 2011; Kodjamanis & Angelopoulos, 2013; Koh & Kim, 2003; Kuhn & McPartland, 1954; Peluchette, Karl, & Fertig, 2013; Wasko & Faraj, 2005; Zhang & Leung, 2015), and by grouping the related questions, we created distinctive attitudinal measures (Table A1, Appendix A). To maximize the validity and minimize the bias of our survey, a pilot phase preceded the data collection, which included $n = 66$ participants. The findings of our pilot phase (Appendix B) show that users tend to use different SNS to address different interests and to separate their personal and professional lives. They tend to use multiple accounts on the same SNS for the same
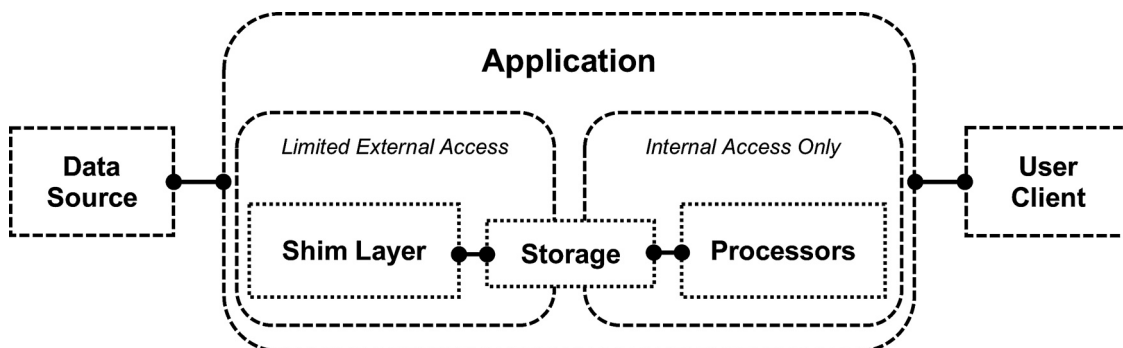

Fig. 3. Application Architecture for SNS.

Table 1
Survey participants' demographics.

Age Group	N (%)
18 − 21 years	3 (1.1)
22 − 25 years	41 (15.2)
26 − 30 years	70 (26.0)
31 − 40 years	108 (40.1)
41 − 50 years	32 (11.9)
51 − 60 years	14 (5.2)
61 − 70 years	1 (0.4)
Gender	**N (%)**
Female	97 (36.1)
Male	172 (63.9)
Highest Level of Education	**N (%)**
High school or equivalent	9 (3.3)
Vocational school (2 year)	11 (4.1)
Some college	45 (16.7)
Bachelor's degree	59 (21.9)
Master's degree	94 (34.9)
Doctoral degree	43 (16.0)
Professional degree	6 (2.2)
Other	2 (0.7)

Table 2
Spread of multiple account use across SNS and standardized chi2 residuals.

SNS	One Account		Multiple Accounts	
	n	Residual	n	Residual
Facebook	202	−1.07	**50**	**2.70***
Twitter	119	−1.08	**33**	**2.72***
LinkedIn	175	1.78	**2**	**−4.49***
Other	14	0.55	0	−1.37

*major contributors to chi^2 (absolute value > 2) highlighted in bold.

Table 3
Binary Logistic Regression Analysis of Use of Multiple SNS Accounts.

	B	S.E.	Exp(B)	95 % C.I. for EXP(B)	
				Lower	Upper
Total number of SNS used	0.39***	0.15	1.47	1.10	1.97
Offline Communication across SNS	0.03	0.05	1.03	0.93	1.13
Breath of use across SNS	0.09*	0.04	1.09	1.01	1.19
Sense of community across SNS	−0.02	0.06	0.98	0.88	1.09
Revelation across SNS	0.04	0.11	1.04	0.84	1.28
Tendency to manage audiences outside of SNS	0.24*	0.12	1.26	1.00	1.60
Tendency to manage audiences inside of SNS	0.01	0.05	1.01	0.91	1.13
General engagement with SNS	0.04	0.06	1.04	0.93	1.16
General satisfaction with SNS	−0.24	0.18	0.79	0.55	1.12
Constant	−3.98	1.05			
Model X2	23.17**				
Nagelkerke R2	0.13				
−2LL	266.96				

reasons. They tend, however, to use the same account on specific SNS to contact specific users, and they also use the same email account to communicate with specific users. The users tend to post the same information on their accounts on the various SNS, and overall, they are satisfied with the use of multiple accounts on the various SNS to self-manage their identities. The findings from our survey, described in greater detail below, enabled us to develop a more nuanced architecture for the proposed processors of our design.

The data collection lasted ten days, and overall 324 SNS users participated. All the participants were native English speakers. Out of these participants, 51 did not complete all the questions and were excluded from the analysis. From the remaining 273 participants, 4 refused the privacy statement and were also excluded from further analysis, leaving a total of 269 participants. The survey was conducted online and addressed an international audience. Demographic information about age, gender, and education was collected. In Table 1, we present the demographic descriptive statistics of the participants in our survey.

As the purpose of the survey was to determine the wider privacy-related needs of SNS users, a cross-sectional analysis related to the nationality of the participants is beyond the scope of this paper. Out of the 269 participants in our study, 251 (93 %) report having a Facebook account, 151 (56 %) a Twitter account, 177 (65 %) a LinkedIn account, and 81 (30 %) an account on other SNS. Facebook, Twitter, and LinkedIn are the most prominent SNS, which is consistent with the results of our pilot study. In terms of multiple account holders, 49 participants (18 %) report having multiple Facebook accounts, 33 multiple Twitter accounts (12 %), 2 multiple LinkedIn accounts (less than 1%), and 10 (approximately 4%) multiple accounts in other SNS, whilst 189 participants (70 %) report not having multiple accounts in any SNS. To ascertain the overall effect between the various SNS and the use of multiple accounts, these differences are explored using a 2*5 chi-squared, followed by a *post hoc* examination of standardized adjusted residuals to identify areas which made a significant contribution to the chi-squared results. We use a binary logistic regression between the use of SNS and our attitudinal measures to explore their relationship. A chi-squared test reveals differences between the use of single and multiple accounts with specific SNS in this context to be significant ($chi^2 = 43.53$, $df = 4$, $p = 8.028*10^{-9}$). The results of the analysis of chi-squared standardized residuals reveal that the use of multiple accounts within the same SNS is more common amongst Facebook and

Twitter users, and considerably less common amongst the users of the other SNS (see Table 2).

Our findings show that participants who tend to use multiple SNS ($Exp[\beta] = 1.47$, CI 1.10–1.97), use them for a wider variety of purposes ($Exp[\beta] = 1.01$, CI 1.01–1.02), and they tend to engage in a higher level of offline audience management ($Exp[\beta] = 1.27$, CI 1.00–1.60). Our results show that the total number of accounts on SNS is the strongest predictor of the multiple uses of SNS (see Table 3), which means that the more SNS a user has at least one account on, the more likely it is for this user to have more than one account on at least one of those SNS. Our analysis reveals differences regarding the information sharing patterns of users on the various SNS. In terms of the total information revelation, a repeated measure MANOVA using Wilks' Lambda reveals significant and large effect amongst the various SNS ($F = 32.009$, Hyp $df = 3$, Error $df = 83$, $p = 1.76*10^{-13}$, $Eta^2 = 0.536$). Moreover, we identify that there is considerable variation in the type of information shared by users on each SNS (Table 2).

A chi-squared analysis demonstrates that these differences are statistically significant ($chi^2 = 330.12$, $df = 27$, $p = 6.87*10^{-54}$), and an examination of the standardized residuals reveals the effects contributing considerably to the chi-square results. To explore the differences in attitudes between the users holding multiple accounts and single accounts, we performed an independent subject MANOVA including the type of SNS user as the independent variable and the attitudinal measures as the dependent variables. The multivariate test was significant ($\lambda = 0.913$, $F = 2.415$, Hyp $df = 10$, Error $df = 253$, $sig = 0.009$), and *post hoc* testing using Sidak correction for multiple tests shows that the participants in our study holding multiple accounts within the same SNS tend to have accounts on more SNS, which they also tend to use for a wider breadth of activities, they tend to reveal more information about themselves, and finally, they are prone to incorporating offline audience-management practices (Table 4).

Table 4

Users of Multiple Accounts (UMA) and Users of Single Accounts (USA).

Dependent Variable	UMA mean	USA mean	Std. Error	Df	F	Sig.	Observed Power
Total number of SNS used	**3.476**	**3.015**	**0.172**	**1 / 262**	**7.175**	**.008***	–
Average Offline Communication across SNS	7.131	6.263	0.508	1 / 262	2.922	.089	.399
Average Breath of Use across SNS	**11.335**	**9.913**	**0.650**	**1 / 262**	**4.789**	**.030***	–
Average Sense of Community across SNS	9.624	9.9624	0.468	1 / 262	0.058	.811	.057
Average Revelation across SNS	**4.254**	**3.802**	**0.226**	**1 / 262**	**3.979**	**.047***	–
Tendency to manage audiences outside of SNS	**9.905**	**9.299**	**0.195**	**1 / 262**	**6.688**	**.010***	–
Tendency to manage audiences inside of SNS	9.905	9.299	0.442	1 / 262	1.880	.172	.277
General engagement with SNS	9.688	8.967	0.487	1 / 262	2.188	.140	.314
General satisfaction with SNS	3.818	3.720	0.134	1 / 262	0.539	.463	.113
Satisfaction with interaction with contacts in SNS	3.800	3.865	0.141	1 / 262	0.210	.647	.074

*Significant effects at $p < 0.05$ highlighted in bold.

Based on our findings, we can conclude that the existing SNS platforms are limited in their capacity to deal with the complex users' needs for two reasons: i) they have been designed for a limited scope, and thus are use-specific, and ii) they do not allow the use of multiple accounts, and thus users have to create accounts on multiple SNS to express multiple identities.

Consequently, the inherent weakness of dominant SNS is that they try to satisfy some of the needs of many users. SNS users, however, can have multiple identities, some of which might be associated with pseudonymous accounts. The importance of controlling privacy and the ability to choose distinct levels of exposure for different combinations of personal data, social contexts, and affordances of online platforms, is reflected in the strategies deployed by the participants of our study. As the findings of our survey show, the attempts to support this richness via the implementation of solutions such as "groups" or "lists" have been proved largely inadequate since the cognitive effort required can result in their misuse or non-use.

4.2. Processor architecture

Following the approach of Peffers et al. (2007), we enhance our conceptual design by proposing two processors that at based on the findings of our survey, which show that users tend to maintain accounts on various SNS, with overlapping groups of contacts. Thus, when users send the same message through their various accounts, their contacts might receive it several times on the various SNS. Our first proposed processor, thus, is an adaptive message filter (Fig. 4, left), which can detect such messages and ensure that the contacts of the user will receive the message only once. Similarly, when users send personal messages, the filter can route them via the SNS where the receiver will most likely respond, by conducting real-time network analytics based on the past communication stored in the storage layer. To illustrate its utility, let us take as example user A who is connected to user B on SNS1 and SNS2. When user A posts the same message on both SNS, user B will receive it through the client just once.

The second processor (Fig. 4, right) protects users from posting copyrightable content, and from the consequences that follow, enabling them to choose the rights they permit. The ease of posting content that all SNS provide, may inhibit its deliberative consideration and lead to

publishing copyrightable content. The processor can identify problematic situations, and warn the users by observing the ToS in the storage layer the rights that need to be transferred. To illustrate its utility, let us take as example user A, who attempts to compliment a post with a picture that has copyrights protection. In such an attempt, the client will promptly warn the user that this picture should not be posted. The conceptual design can be strengthened with 3rd party processors, as it is built on open standards for this purpose (e.g. Boehmke, Hazen, Boone, & Robinson, 2020).

4.3. Architecture communication

For the communication of the proposed conceptual design with the current and future SNS, we use existing secure communication protocols, which can ensure its security and efficiency. Similarly to the use of multiple online identities on SNS, people use multiple email addresses to separate their communication based on their personal and professional identities; having personal email addresses for ease of use and longevity, as well as professional ones for their relevant professional communication. These identities and respective email addresses exist together only within the private context of users' devices.

There are two prominent communication protocols for sending and receiving emails. The most prominent communication guidelines enabling applications to send emails is the *simple mail transfer protocol* (SMTP), while for retrieving and storing emails the most prominent ones are the *Internet message access protocol* (IMAP), and the *post office protocol 3* (POP3), with the former enabling users to synchronize emails across multiple devices. The use of SMTP and IMAP enables the incorporation of existing protocols to transmit and store information. Concurrently, the *multi-purpose Internet mail extensions* (MIME) allow the exchange of many data types as well as metadata related to addressing and provenance. A wide range of existing clients can understand MIME, enabling the experimentation with storing SNS data as MIME, and displaying them without format changes, enabling the support of multi-party encryption.

Such an approach can address format standardization and provide a flexible transport-independent asymmetric encryption for per-user and per-service privacy and authentication. Users' multiple identities are linked only on the user-side based on out-of-band information
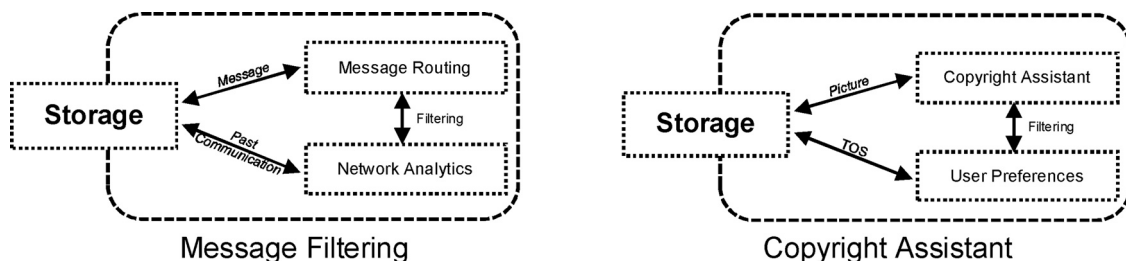

Fig. 4. Proposed Processor Architecture.

indicating that they belong to a specific SNS user. When identities belong to several pods, they can be blurred and have a semantic meaning for the users who can resolve them. Such an approach enables SNS communication through key management with authentication support for messaging, allowing trusted clients to interpret incoming messages, which can become trusted via face-to-face interactions or out-of-band mechanisms.

5. Discussion

The proposed conceptual design redresses the asymmetric distribution of power between users and online platforms, by being *open*, and self-hosted on the user-side. It, thus, enables online platforms to provide agile and lightweight applications that are relevant to real-time individual contexts, and able to run on the user-side, which only return the outcomes of analysis (Fig. 1). Each application installed on the user-side for an online platform, also generates data in the form of analyses results that can be stored and used as input by the applications of other online platforms, and thus, the consumption of data on the user-side leads to the production of new data about the user, which always remain stored on the user-side.

As shown by our survey participants, the limitations of online platforms have compelled users to develop complicated and cumbersome strategies to manage their identities and privacy. In the literature and popular media, the discussion about users' privacy is becoming increasingly relevant and heated, due to the consequences that follow when the stewardship of personal data is compromised, and there have been many attempts to provide solutions to this problem (Perera et al., 2017). To date, however, most of such proposed solutions are software-based and cloud-hosted, leading to another set of problems related to the security, performance, and interactivity of the data (Mortier et al., 2016). However, that said, such software-based and cloud-hosted solutions can run as processors in the proposed conceptual design we describe here, minimizing in this way the limitations they present for users' privacy.

5.1. Implications

Our work makes an artefactual contribution (Ågerfalk & Karlsson, 2020) towards resolving the privacy and provenance issues highlighted in the IM research agenda (Abbasi et al., 2016; de Camargo Fiorini et al., 2018; Ghasemaghaei, 2019, 2020). The proposal that personal data reside with the user can provide greater control over who can access the data. Such a change in the paradigm of the *instrumental* ways that personal data are stored and used (Chang et al., 2014; Kitchin, 2014) can catalyse a change to "the entire social theory that goes with them" (Latour, 2009, p. 155).

The proposed conceptual design integrates message filtering, ensuring that SNS users receive each message once. Currently, it is not uncommon for SNS users to receive a specific message several times from another user since many users, out of convenience, interconnect their accounts on the various SNS and distribute the same material on them. Moreover, our conceptual design provides intelligent lists for sharing messages, enabling users to distribute certain material only to those that it is relevant to. Exercising control in this way can be extremely valuable, as, for instance, the family members of a user might not have an interest in receiving updates concerning a hobby, or it might be embarrassing—or worse—to share personal photos with colleagues. Our conceptual design enables an intelligent application, which takes into consideration the various SNS distribution channels as well as the multiple identities of users, enabling suggestions regarding the possible recipients of a message, taking into account the content of the message, the interests of the users, and the properties of the routes available to reach all the recipients, along with information regarding the contacts that should not receive it.

Implications can also be distilled from the results of our survey,

since our findings demonstrate that it is necessary to revisit the way that we design SNS, to provide users with better ways to self-manage their multiple context-dependent online identities. This can provide users with greater utility, creativity, and flexibility in the use of their social networks along with control over such use. The self-management of multiple context-dependent online identities is still an issue to be pursued by both academia and the industry, and there is a profound need for better tools to be implemented either by the SNS or by third-parties that will take up the challenge of implementing new tools tailored to the identified needs of SNS users, since the tools that are already available and provided by the existing SNS for the self-management of multiple context-dependent online identities are insufficient and often neglected by the users.

Our study, thus, contributes to the extant IM literature by motivating and implementing a conceptual design in response to the challenges related to the stewardship of personal data on online platforms in general, and in SNS more specifically, without compromising users' privacy (Agarwal & Dhar, 2014; Goes, 2014). Moreover, our work contributes to the IM discussions on privacy-focused SNS, by applying the proposed conceptual design to the context of SNS, to address issues related to privacy and identity (Krasnova et al., 2017), and to pave the path to a broader set of possible applications. Our conceptual design does not intend to provide a "better" alternative to existing systems in use, rather one that is privacy-focused, and can enable opportunities for online platforms with privacy by design (Bu, Wang, Jiang, & Liang, 2020), as well as give rise to theory building in line with the IM research agenda (Kar & Dwivedi, 2020). Our work, thus, does not attempt to *close* the discussions on the topic, rather *open* it to novel avenues of research and development.

The proposed conceptual design also bears practical implications for the management of online platforms. Specifically, organizations can be relieved of the burden of maintaining data, and the data will not be decoupled from information about their provenance and context of origin. As personal data reside with on the user-side, online platforms do not need costly infrastructures to retain them. Moreover, the proposed conceptual design can enable online platforms to provide applications relevant to real-time individual contexts, giving rise thus to novel business practices for personal data use and storage, and spark novel business models related to leveraging the user-side affordances. For instance, using the outcome of the analysis of an application as input for another application can give rise to new communities of 3rd party developers that will be able to provide context for the input of their applications, with users able to define the level of access that such applications would have to their data (Price et al., 2015).

5.2. Limitations and future research

Certain limitations need to be acknowledged, emerging mainly from a series of real-world difficulties that surround the implementation and use of our proposed conceptual design by both users and SNS organizations (e.g. Chen, Lu, & Tang, 2019; Lu, Li, Ioannou, & Tussyadiah, 2019). The main limitation of our work here is the lack of a prototype implementation and real-world evaluation (Snyder, 2003) of the proposed conceptual design and its application on SNS, as it would be prescribed by design science (Peffers et al., 2007). This phase would require the collaboration of existing SNS in implementing our conceptual design, along with the possible business models that could complement it, which was beyond the scope of our study. Our work, however, can provide the bedrock for future research on the topic, by enabling various prototyping endeavours of our framework as well as studies on the possibilities for its adoption by both the organizations behind popular SNS, as well as the users. On this note, our proposal is founded on the assumption that users will be willing to be responsible for the maintenance of their data. Such a view is optimistic, since users are known to resist changes in the way they do things (Ali, Zhou, Miller, & Ieromonachou, 2016; Chen et al., 2019; Lapointe & Rivard, 2005;

Sykes, Venkatesh, & Johnson, 2014; Venkatesh, Morris, Davis, & Davis, 2003). Such a limitation, however, can spark new business opportunities, for instance, on providing users with personal data storage, and redundancy solutions. Furthermore, this can lead developers to create novel, secure, reliable, and cost-effective applications. Future research should investigate the possibilities for the adoption of such a conceptual design by users, and explore their willingness to take responsibility for their data to regain their privacy.

Another limitation emerges from potential resistance against the adoption of our conceptual design by the existing SNS, which tend to profit via targeted advertising generated through the collection and analysis of personal data. Our design could be perceived as an impediment, negating direct access to personal data, and consequently, to profiting. Our conceptual design, however, primarily addresses the growing public pressure and advocacy for personal data to be protected. While the aforementioned reason might inhibit the explicit implementation of our proposed conceptual design by the existing SNS, it can give rise to new business ventures and applications through APIs for implementation from 3rd parties. As the investigation of such opportunities would go beyond the scope of our paper, we would encourage future research endeavours to explore such business models based on our proposal.

One more limitation emerges from the application of the conceptual design on SNS. The message remains the focus of SNS use, be it text, images or videos. This means that data in various formats are transferred through the Internet from a sending user to a group of defined recipients. Each one of the existing SNS uses a different structure for this purpose, and in many cases provide a different level of richness of the message, with extensive metadata. A critical task for the proposed conceptual design, thus, is the creation of maps for the various SNS message structures. While this provides interoperability to our conceptual, it poses a limitation for its implementation. There is, however, underlying flexibility which is valuable: if the initially chosen mapping proves to be unsuitable in some way, it is a simple matter for those concerned to create another mapping, and even for multiple mappings to be operational in parallel. The flexibility engendered by avoiding "locking in" the system to a single ontology that attempts to describe all formats and uses of data, allows the system to be flexible and to respond to changes in context and demand, enabling it to be continually updated and extended to support new uses. We, thus, encourage future research endeavours to further extend the proposed conceptual design with the mapping of existing SNS message structures, and evaluate their usability.

Finally, here we used a design science approach to develop a more nuanced conceptual design that addresses specific user needs; future research, however, could incorporate a canonical action research perspective (Davison, Martinsons, & Ou, 2012; Davison, Martinsons, & Kock, 2004), or an action design research perspective (Sein, Henfridsson, Purao, Rossi, & Lindgren, 2011) to empirically investigate the topic within SNS organizations, and implement designs that could

Table A1
Attitudinal measures and questions from the questionnaire.

Attitudinal Measure	Questions
Offline Communication across SNS (Koh & Kim, 2003)	Do you chat with your contacts?
	Do you call your contacts from this SNS on the phone?
	Do you meet your contacts from this SNS in offline settings?
	Do you exchange SMS messages with your contacts from this SNS?
	Do you exchange e-mails with your contacts from this SNS?
	Do you use VoIP to communicate with your contacts from this SNS?
Breath of use across SNS (Dholakia et al., 2004, Kodjamanis & Angelopoulos, 2013)	How often do you use (SNS list) to satisfy the following needs?
	• To contribute / get information
	• To stay in touch with friends / family
	• To stay in touch with colleagues
	• To impress friends / family
	• To impress colleagues
	• To play / relax
	• To search for jobs
	• To share pictures
Sense of community across SNS (Gefen & Straub, 2004)	There is a sense of human contact (SNS list)
	There is a sense of personalness (SNS list)
	There is a sense of sociability (SNS list)
	There is a sense of human warmth (SNS list)
	There is a sense of human sensitivity (SNS list)
Revelation across SNS (Duffy, Shaw, & Stark, 2000)	Do you include your phone number(s) on your profile(s)?
	Do you include your home address on your profile(s)?
	Do you include information about your interests on your profile(s)?
	Do you include information about your personality on your profile(s)?
	Do you spend time personalizing your profile page(s)?
	Do you use your real name on your profile page(s)?
	Do you allow everyone to view your profile(s)?
Tendency to manage audiences outside of SNS (Koh & Kim, 2003)	I have more than one email account
	I use different email accounts to address different interests
	I use the same email account to communicate with certain contacts
	I have more than one phone number
	I use different phone numbers to address different interests
	I use the same phone number to communicate with certain contacts
Tendency to manage audiences inside of SNS (Koh & Kim, 2003)	I use different accounts to address different interests
	I use different accounts to separate personal / professional
	I use a single account to communicate with contacts from a certain SNS
	I post the same information on all my different accounts on an SNS
General engagement with SNS (Koh & Kim, 2003, Wasko & Faraj, 2005)	I participate actively on (SNS list)
	I reply to my contacts who need help on (SNS list)
	I am pleased to interact with my contacts on (SNS list)
General satisfaction with SNS (Duffy et al., 2000)	I am pleased with my experience on (SNS list)

Table B1

Descriptive Statistics of Variables (n = 66).

	Percentage	Members	Std. Dev
One Account	60.6 %	40	0.492
Multiple accounts	19.7 %	13	0.400
Different SNS for different interests	45.4 %	30	0.501
Different SNS for personal and professional	43.9 %	29	0.500
Different accounts for different interests	16.7 %	11	0.375
Different accounts for personal and professional	13.6 %	9	0.345
More than one email account	54.5 %	36	0.501
Different email to address different interests	48.5 %	32	0.503

Table B2

Correlation Matrix.

	1	2	3	4	5	6	7	8
1. One Account	1							
2. Multiple accounts	0.399	1						
3. Different SNS for different interests	0.736	0.466	1					
4. Different SNS for personal and professional	0.714	0.559	0.786	1				
5. Different accounts for different interests	0.361	0.903	0.408	0.505	1			
6. Different accounts for personal and professional	0.320	0.802	0.347	0.449	0.888	1		
7. More than one email account	0.883	0.376	0.711	0.686	0.327	0.274	1	
8. Different email to address different interests	0.782	0.358	0.576	0.607	0.379	0.321	0.886	1

Table B3

Results of the analysis of variance.

	MS	F	Prob > F
Different SNS for different interests	2.26717172	17.76	0.0001
Different SNS for personal and professional	3.26698015	29.15	0.0000
Different accounts for different interests	8.51212121	282.67	0.0000
Different accounts for personal and professional	6.72009569	115.64	0.0000
More than one email account	1.47272727	10.51	0.0019
Different email to address different interests	1.338291	9.41	0.0032

overcome the specific limitations we report in this study.

6. Conclusions

Our work focuses on the issues and challenges related to personal data and proposes a conceptual design that makes use of the principles of Distributed Computing and SoC, which can resolve the challenges of integrity, storage, and analysis that relate to users' privacy. Our conceptual design addresses the asymmetrical distribution of power between users and online platforms by shifting from a user-platform relationship to one in which access to personal data is controlled by the user. For organizations, the benefits of our conceptual design include being liberated from the burden of maintaining and updating large centralized data stores and having access to relevant high-granularity data from the entire distributed population of user-held data, which are not decoupled from their context of origin. We demonstrate the affordances of our conceptual design by illustrating its application on SNS, where we describe an architecture that addresses issues relating to the privacy, and identity of users and opens up SNS to richer applications while referring to the limitations that it might present. Drawing on insights about the behaviours and strategies deployed by SNS users, our conceptual design can enable greater utility, flexibility, and creativity in the use of online platforms, and provide users with greater control over their data. Being a design science study that leads to a novel conceptual design, our work is important and timely for IM research and practice, and it gives rise to two main directions for future research

that can provide opportunities for online platforms with privacy by design (Bu et al., 2020). Firstly, there is scope for evaluating our conceptual design and comparing it to the current models used by most online platforms that exploit personal data. Secondly, there is the prospect of applying our conceptual design in areas beyond SNS such as in banking, energy use, commerce, governance, and healthcare. These avenues for future research will enable theory building in line with the IM research agenda (Kar & Dwivedi, 2020).

CRediT authorship contribution statement

Spyros Angelopoulos: Conceptualization, Methodology, Formal analysis, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Michael Brown:** Formal analysis, Validation, Writing - review & editing. **Derek McAuley:** Conceptualization, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Yasmin Merali:** Writing - review & editing, Supervision. **Richard Mortier:** Conceptualization, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Dominic Price:** Software, Writing - original draft, Visualization.

Declaration of Competing Interest

The authors report no declarations of interest.

Appendix A

Appendix B. Analysis of pilot phase

Tables B1, B2, B3

References

Abbasi, A., Sarker, S., & Chiang, R. H. (2016). Big data research in information systems: Toward an inclusive research agenda. *Journal of the Association for Information Systems, 17*(2), i–xxxii.

Agarwal, R., & Dhar, V. (2014). Editorial: Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443–448.

Ågerfalk, P. J., & Karlsson, F. (2020). Artefactual and empirical contributions in information systems research. *European Journal of Information Systems, 29*(2), 109–113.

Ali, M., Zhou, L., Miller, L., & Ieromonachou, P. (2016). User resistance in IT: A literature review. *International Journal of Information Management, 36*(1), 35–43.

Angelopoulos, S., & Merali, Y. (2015). Bridging the divide between virtual and embodied spaces: Exploring the effect of offline interactions on the sociability of participants of topic-specific online communities. *Proceedings of 48th HICSS Conference*.

Angelopoulos, S., & Merali, Y. (2017). *Sometimes a cigar is not just a cigar: Unfolding the transcendence of boundaries across the digital and physical.* Seoul, Korea: ICIS.

Bayramusta, M., & Nasir, V. A. (2016). A fad or future of IT?: A comprehensive literature review on the cloud computing research. *International Journal of Information Management, 36*(4), 635–644.

Boehmke, B., Hazen, B., Boone, C. A., & Robinson, J. L. (2020). A data science and open source software approach to analytics for strategic sourcing. *International Journal of Information Management, 54*, Article 102167.

Bu, F., Wang, N., Jiang, B., & Liang, H. (2020). "Privacy by Design" implementation: Information system engineers' perspective. *International Journal of Information Management, 53*, Article 102124.

Chang, R. M., Kauffman, R. J., & Kwon, Y. (2014). Understanding the paradigm shift to computational social science in the presence of big data. *Decision Support Systems, 63*, 67–80.

Chen, Q., Lu, Y., & Tang, Q. (2019). Why do users resist service organization's brand mobile apps? The force of barriers versus cross-channel synergy. *International Journal of Information Management, 47*, 274–282.

Conboy, K. (2009). Agility from first principles: Reconstructing the concept of agility in information systems development. *Information Systems Research, 20*(3), 329–354.

Davison, R. M., Martinsons, M. G., & Ou, C. X. (2012). The roles of theory in canonical action research. *MIS Quarterly, 36*(3), 763–786.

Davison, R., Martinsons, M. G., & Kock, N. (2004). Principles of canonical action research. *Information Systems Journal, 14*(1), 65–86.

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.

Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network-and small-group-based virtual communities. *International Journal of Research in Marketing, 21*(3), 241–263.

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61–80.

Duffy, M. K., Shaw, J. D., & Stark, E. M. (2000). Performance and satisfaction in conflicted interdependent groups: When and how does self-esteem make a difference? *Academy of Management Journal, 43*(4), 722–783.

Ellison, N., Heino, R., & Gibbs, J. (2006). Managing impressions online: Self-presentation processes in the online dating environment. *Journal of Computer-Mediated Communication, 11*(2), 415–441.

Gefen, D., & Straub, D. W. (2004). Consumer trust in B2C e-commerce and the importance of social presence. Experiments in e-products and e-services. *OMEGA, 32*(6), 407–424.

Georgiadou, E., Angelopoulos, S., & Drake, H. (2019). Big data analytics and international negotiations: Sentiment analysis of Brexit negotiating outcomes. *International Journal of Information Management* Article 102048.

Ghasemaghaei, M. (2019). Understanding the impact of big data on firm performance: The necessity of conceptually differentiating among big data characteristics. *International Journal of Information Management* Article 102055.

Ghasemaghaei, M. (2020). The role of positive and negative valence factors on the impact of bigness of data on big data analytics usage. *International Journal of Information Management, 50*, 395–404.

Goes, P. (2014). Big data and IS research. *MIS Quarterly, 38*(3), iii–vi.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly, 37*(2), 337–355.

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2020). The role of artificial intelligence and data network effects for creating user value. *The Academy of Management Review*. https://doi.org/10.5465/amr.2019.0178.

Grønli, T. M., Ghinea, G., & Bygstad, B. (2013). Exploring solutions for mobile companionship: A design research approach to context-aware management. *International Journal of Information Management, 33*(1), 227–234.

Guo, Y., Lu, Z., Kuang, H., & Wang, C. (2020). Information avoidance behaviour on social network sites: Information irrelevance, overload, and the moderating role of time pressure. *International Journal of Information Management* Article 102067.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly, 28*(1), 75–105.

Holmström, J., Ketokivi, M., & Hameri, A. P. (2009). Bridging practice and theory: A design science approach. *Decision Sciences, 40*(1), 65–87.

Hürsch, W. L., & Lopes, C. V. (1995). *Separation of concerns* Technical report. North Eastern University.

Ioannou, A., Tussyadiah, I., & Lu, Y. (2020). Privacy concerns and disclosure of biometric and behavioural data for travel. *International Journal of Information Management, 54*, Article 102122.

Janse, N., Ou, C. X., Angelopoulos, S., Davison, R. M., & Jia, J. W. (2017). Do security breaches matter to consumers? *the Proceedings of ICEB, 50* https://aisel.aisnet.org/iceb2017/50.

Kar, A., & 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.

Karl, K. A., & Peluchette, J. V. E. (2011). Friending professors, parents and bosses: A Facebook connection conundrum. *The Journal of Education for Business, 86*(4), 214–222.

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society, 1*(1), 1–12.

Kodjamanis, A., & Angelopoulos, S. (2013). Consumer perception and attitude towards advertising on social networking sites: The case of Facebook. *International Conference of Communication, Media Studies and Design*. Available online at: http://goo.gl/9pVlAv.

Koh, J., & Kim, Y. G. (2003). Sense of virtual community: Determinants and the moderating role of virtual community origin. *International Journal of Electronic Commerce, 8*(2), 75–93.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788–8790.

Krasnova, H., Veltri, N. F., Eling, N., & Buxmann, P. (2017). Why men and women continue to use social networking sites: The role of gender differences. *The Journal of Strategic Information Systems, 26*(4), 261–284.

Kuhn, M. H., & McPartland, T. S. (1954). An empirical investigation of self-attitudes. *American Sociological Review, 19*(1), 68–76.

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly, 29*(3), 461–491.

Latour, B. (2009). Tarde's idea of quantification. In M. Candea (Ed.). *The social after Gabriel Tarde: Debates and assessments* (pp. 145–162). London: Routledge.

Liu, Y., Soroka, A., Han, L., Jian, J., & Tang, M. (2019). Cloud-based big data analytics for customer insight-driven design innovation in SMEs. *International Journal of Information Management* Article 102034.

Lowry, P. B., Gaskin, J. E., & Moody, G. D. (2015). Proposing the multimotive information systems continuance model (MISC) to better explain end-user system evaluations and continuance intentions. *Journal of the Association for Information Systems, 16*(7), 515–579.

Lu, Y., Li, S., Ioannou, A., & Tussyadiah, I. (2019). From data disclosure to privacy nudges: A privacy-aware and user-centric personal data management framework. *Proceedings of the International Conference on Dependability in Sensor, Cloud, and Big Data Systems and Applications, 262*–276.

Lu, Y., Ou, C., & Angelopoulos, S. (2018). Exploring the effect of monetary incentives on user behaviour in online sharing platforms. *Proceedings of 51st HICSS Conference*, 3437–3444.

Lycett, M. (2013). "Datafication": Making sense of (big) data in a complex world. *European Journal of Information Systems, 22*(4), 381–386.

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research, 15*(4), 336–355.

Montealegre, R., Hovorka, D., & Germonprez, M. (2014). A coevolutionary view of information services development: lessons from the US National Oceanic and Atmospheric Administration. *Journal of the Association for Information Systems, 15*(9), 577–613.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D., Crowcroft, J., & Crabtree, A. (2016). *Human-data interaction. The encyclopedia of human-computer interaction.* Interaction Design Foundation ISBN 9788792964007.

Nunamaker, J. F., Jr., Chen, M., & Purdin, T. (1991). Systems development in information systems research. *Journal of Management Information Systems, 7*(3), 89–106.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77.

Peluchette, J. V. E., Karl, K., & Fertig, J. A. (2013). Facebook 'friend' request from the

boss: Too close for comfort? *Business Horizons, 56*, 291–300.

Perera, C., Wakenshaw, S. Y., Baarslag, T., Haddadi, H., Bandara, A. K., Mortier, R., et al. (2017). Valorising the IoT databox: Creating value for everyone. *Transactions on Emerging Telecommunications Technologies, 28*(1), e3125.

Price, D., McAuley, D., Mortier, R., Greenhalgh, C., Brown, M., & Angelopoulos, S. (2015). Inter-social-networking: Accounting for multiple identities. *The Proceedings of the International Conference on Social Computing and Social Media* (pp. 242–252).

Schwämmlein, E., & Wodzicki, K. (2012). What to tell about me? Self-presentation in online communities. *Journal of Computer-Mediated Communication, 17*(4), 387–407.

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly, 35*(1), 37–56.

Senyo, P. K., Addae, E., & Boateng, R. (2018). Cloud computing research: A review of research themes, frameworks, methods and future research directions. *International Journal of Information Management, 38*(1), 128–139.

Shibchurn, J., & Yan, X. (2015). Information disclosure on social networking sites: An intrinsic–extrinsic motivation perspective. *Computers in Human Behavior, 44*, 103–117.

Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces.* Morgan Kaufmann.

Spanaki, K., Gürgüç, Z., Mulligan, C., & Lupu, E. (2019). Organizational cloud security and control: A proactive approach. *Information Technology and People, 32*(3), 516–537.

Spanaki, K., Karafili, E., Sivarajah, U., Despoudi, S., & Irani, Z. (2020). Artificial intelligence and food security: swarm intelligence of AgriTech drones for smart AgriFood operations. *Production Planning and Control* in press, ISSN: 0953-7287.

Sullivan, Y. W., & Koh, C. E. (2019). Social media enablers and inhibitors: Understanding their relationships in a social networking site context. *International Journal of Information Management, 49*, 170–189.

Sykes, T. A., Venkatesh, V., & Johnson, J. L. (2014). Enterprise system implementation and employee job performance: Understanding the role of advice networks. *MIS Quarterly, 38*(1), 51–72.

Venkatesh, V., Morris, M., Davis, G., & Davis, F. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478.

Wang, L., Yan, J., Lin, J., & Cui, W. (2017). Let the users tell the truth: Self-disclosure intention and self-disclosure honesty in mobile social networking. *International Journal of Information Management, 37*(1), 1428–1440.

Wang, T., Duong, T. D., & Chen, C. C. (2016). Intention to disclose personal information via mobile applications: A privacy calculus perspective. *International Journal of Information Management, 36*(4), 531–542.

Wasko, M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly, 29*(1), 35–57.

Zhang, Y., & Leung, L. (2015). A review of social networking service (SNS) research in communication journals from 2006 to 2011. *New Media & Society, 17*(7), 1007–1024.

Zuo, M., Angelopoulos, S., Ou, C. X., Liu, H., & Liang, Z. (2020). *Identifying dynamic competition in online marketplaces through consumers' clickstream data.* Available at SSRN 3598889.