Journal of Strategic Information Systems 26 (2017) 185–190

Contents lists available at ScienceDirect

Journal of Strategic Information Systems

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

Datification and its human, organizational and societal effects: The strategic opportunities and challenges of algorithmic decision-making

R.D. Galliers[a,*], S. Newell[b], G. Shanks[c], H. Topi[d]

[a] Bentley University, USA and Loughborough University, UK
[b] Sussex University, UK
[c] University of Melbourne, Australia
[d] Bentley University, USA

This special section on datification[1] attempts to go some way to answer the call for research proposed by Newell and Marabelli (2015) given the increasing reliance being placed on algorithmic decision-making in modern society. In calling for papers on this topic we aimed to critically assess society's apparent taken-for-granted and unknowing acquiescence to this increasingly prevalent phenomenon (Markus and Topi, 2015).

As we noted in the call for papers (Galliers et al., 2015, p. II), Newell and Marabelli argued that "… the many digital devices that are increasingly in continuous use are capable of enabling the monitoring of 'the minutiae of an individual's everyday life'. Such data are often processed by pre-determined algorithms that lead to decisions that follow on directly without further human intervention (often with the claim that the decisions are for the individual's benefit)". This is, of course, something that they – and we – question. While the strategic value of such data for organizations can be considerable and is doubtless growing, the implications for individuals and wider society are less clear and debatable. As Newell and Marabelli point out, we most often remain "unaware of how the data they [digital devices] produce are being used, and by whom and with what consequences". The aim of their *JSIS* Viewpoint article was thus to lay a foundation for this discussion to take place – in the IS community and beyond.

We also noted in the call for submissions that vast amounts of "digital trace data are collected through digitized devices (captured, for example, via social networks, online shopping, blogs, ATM withdrawals and the like) and through in-built sensors. As such, they fall under the 'big data' umbrella (Hedman et al., 2013; Wu and Brynjolfsson, 2009)".[2] But we should not discount 'little data' in our discussion. As Newell and Marabelli (2015) pointed out, "While using big data and algorithmic decision-making … this targeting can now be taken further when data are used not to predict group trends but to predict the behavior of a specific individual". Thus, 'little data', is based on 'big data' but has its focus on individuals, using the vast computing capacity that is available today to collect and analyze what is extremely granular data (Munford, 2014) – such as whether an individual is driving safely or not (e.g., Abbas et al., 2014).

In a nutshell, then, a major concern with 'datification' (whether in relation to 'big' or little' data) is that "somebody else may . . . use the data … often with purposes different from those originally intended" (Newell and Marabelli, 2015).

In response to our call, we accepted two papers, by Günther et al. (2017) and by Marjanovic and Cecez-Kecmanovic (2017), and invited a commentary on these papers and the broader issues of 'datification' by Markus (2017). As Markus points out, our contributions are but a "microcosm" of the kind of debate that is occurring in the wider academic literature on datification. Nevertheless, they do at least provide some food for thought and, importantly, give rise to an agenda for future research in this significant, growing,

* Corresponding author.
E-mail addresses: rgalliers@bentley.edu (R.D. Galliers), Sue.Newell@sussex.ac.uk (S. Newell), gshanks@unimelb.edu.au (G. Shanks), htopi@bentley.edu (H. Topi).

[1] Also called 'datafication" (e.g., in Hansen, 2015; Lycett, 2013; Mai, 2016)

[2] As noted in the call for papers, big data analytics is a similar concept to the older and more familiar concept of business intelligence that has been studied for over a decade (e.g., Negash, 2004; Power, 2002; Rouibah and Ould-ali, 2002; Shollo and Galliers, 2015; Thomsen, 2003), with the difference that, in the big data context, the sources and types of data are significantly more varied and often get their relevance from real-time processing.

http://dx.doi.org/10.1016/j.jsis.2017.08.002

but still emerging topic area.

The paper by Günther et al. (2017) nicely sets the scene for us in at least two ways. First, it provides a review of the extant literature on big data analytics, identifying "six debates central to how organizations realize value from big data … and two socio-technical features … that influence value realization: portability and interconnectivity". In addition, the authors identify three different levels of analysis for the IS academy to conduct future research. These are the work-practice, organizational and supra-organizational levels, with two debates being identified for each.

At the *work-practice level*, the debates that Günther et al. identify – we shall call them *tensions* – relate to 1) inductive *versus* deductive approaches to big data analytics, and 2) algorithmic *versus* human-based intelligence. Key issues identified include the collection of data without a predefined purpose that, in their words, promote "a bottom-up approach to big data collection, exploration and analysis". Such inductive approaches as this are meant to lead to the identification of "previously unknown patterns or distinctions". They also note, however, that more deductive, hypothesis-driven approaches "are common … in healthcare settings [for example], where data are collected, processed, and visualized for specific purposes (Tan et al., 2015)". There are risks inherent in either approach, of course, the one being that data may be used out of context, the other relating to confirmation bias. Thus, a more balanced approach is advocated by some (e.g., Bholat, 2015), leading to the broader tension associated with algorithmic and human-based intelligence to which we now turn.

Those who point to the benefits associated with algorithmic intelligence – Masden (2015) and Van der Vlist (2016) among others – cite the emergent nature of such intelligence and the innovative concepts thus derived, while praising the avoidance of preconceptions in this approach. Those who place greater emphasis on human intelligence – Sharma et al. (2014) and Seddon et al. (2017) among others – again express concerns about the 'unknowing' and 'out of context' nature of what might be termed the 'blind' dependence on the algorithmic approach. Whatever the strengths and weaknesses of either stance, however, Günther et al. argue for more research on actual practices (cf. Peppard et al., 2014), especially given the number of conceptual rather than empirical studies that have thus far been conducted. For example, actors in the study conducted by Shollo and Galliers (2015) argue that "data should be supplemented with human experience, common sense, and contextual knowledge that are hard to capture by data [alone]". One of the dangers of over-reliance on algorithms is the potential of relevant tacit knowledge being lost or replaced, as noted by Markus (2015) and Newell and Marabelli (2015). Günther et al. (2017) conclude that, "As of yet, it remains unclear under what particular conditions organizational actors are able to generate insights through inductive or deductive approaches, or a combination of both. Nor is it clear what specific contributions human and algorithmic intelligence add to the creation of insights in different situations (e.g., stable and routine practices versus emergent and temporal situations)."

The tensions identified by Günther et al. at the *organizational level* refer to 3) centralized *versus* decentralized big data capability structures, and 4) business model improvement as against more radical innovations. Capabilities with respect to organizations developing, mobilizing and using both technical and human resources (cf., Peppard and Ward, 2004) have been a subject of considerable attention in *JSIS* over the years (e.g., Huang et al., 2015; Daniel et al., 2014). With regard to big data analytics, "organizations face questions regarding not only how to acquire or develop [these] resources (Brinkhues et al., 2015; Tambe, 2014), but also how to structure them in teams or departments". Arguments for the development of centralized competency centers tend to be focused on the (current) shortage of analytical skills (e.g., Davenport et al., 2010 in Sharma et al., 2014). Counter arguments that highlight concerns about the potential of damaging communication between and limiting involvement with other organizational actors have also been raised. This, in turn, has led to the identification of "synergistic benefits of centralized capability structures … [alongside] specific expertise associated with decentralizing (Sidorova and Torres, 2014)". Importantly, Günther et al. (2017) make the point that literature is still scarce as to what is appropriate and how this may be achieved in practice. While examples of centralized capability structures have appeared, such as in Bholat (2015), "it is often not clear how these structures are put in place, how they interact with business units, or how they produce value". Likewise, little empirical evidence exists to support a more decentralized approach.

Similarly, with regard to the tension between more incremental improvements on the one hand and more radical innovation on the other, Günther et al. see this a rising topic of discussion, citing Loebbecke and Picot (2015) and Woerner and Wixom (2015). Once again, they point to the lack of empirical studies with few cases having been published concerning "improvements in or innovations to business models based on big data (Gartner, 2013)".

At the *supra-organizational* level, Günther et al. (2017) point to the role of relevant stakeholders such as "universities and research institutes (Struijs et al., 2014; Ekbia et al., 2015), governments (Chatfield et al., 2015; Kim et al., 2014), data providers (Greenaway et al., 2015), users, and customers (Constantiou and Kallinikos, 2015; Ekbia et al., 2015; Kennedy and Moss, 2015)". They identify two tensions in terms of "how organizations manage data access, and how they deal with stakeholder interests such as ethical concerns and regulation". The tensions relate to 5) controlled *versus* open big data access, and those between 6) minimizing or simply neglecting the *social* risks of big data value realization. By social risks they mean the potential of (inadvertently) revealing personal, sensitive information, in terms of, for example, "privacy, identity theft, illegal discrimination, unjust classification" (see also Markus, 2015).

In terms of the controlled *versus* open big data access tension, Günther et al. (2017) highlight the literature that points to organizations relying on effective data exchange across their partner network (Malgonde and Bhattacherjee, 2014) and engaging in practices of data disclosure and screening in doing so (Jia et al., 2015). "In this respect, different inter-organizational governance modes have been identified, depending on the extent to which access to data is controlled or open (Van den Broek and Van Veenstra, 2015)." Concerns exist of course, for example, in relation to privacy and security (Chatfield et al., 2015), or the potential negative impacts of sharing proprietary or competitive information that may negatively impact an organization's unique strategic position (e.g., Jagadish et al., 2014; Greenaway et al., 2015) – issues raised previously in *JSIS* by, for example, Marabelli and Newell (2012),