Value Creation from Data – Why is this a BPM Problem?

Shazia Sadiq[1], Ida Asadi Someh[2], Tianwa Chen[1], Marta Indulska[2]

[1] School of Information Technology and Electrical Engineering,
The University of Queensland, Brisbane, Australia
shazia@itee.uq.edu.au, tianwa.chen@uq.edu.au
[2] Business School, The University of Queensland, Brisbane, Australia
i.asadi@business.uq.edu.au, m.indulska@business.uq.edu.au

Abstract. The data deluge and associated technological proliferations have significantly changed the landscape of how businesses are run. These changes, in turn, necessitate profound changes in how business processes are managed. Yet, as organisations aspire towards embracing data-driven approaches both technically and culturally, the socio-technical barriers for value creation from data are becoming increasingly evident. This paper highlights the important role that BPM research and practice can play in lifting those barriers.

Keywords: Value Creation, Data-driven Organization, Technology Adoption, Business Analytics, Data Quality

1 Why is value creation from data so hard?

The technical advancements in data science and machine learning, as well as the third wave of AI [10], have raised expectations of business transformation. However, how organisations exploit and adapt to these advancements remains an open question for now. Extant research from Information Systems provides many insights in the context of value creation from IT assets and capabilities [25], but value creation from data challenges many of those findings. Furthermore, in the current characterisations of big data, i.e., the so-called Vs,, a number of well-established data management practices are no longer valid [38], leaving organisations to face the complex and unstable reality of the data-driven promise. A number of aspects have contributed to these difficulties:

First, data re-purposing [41], has resulted in a distance between the design and use intentions of the data, and is causing a fundamental shift in the way data is managed and used. Traditional data modelling and design principles are thus challenged in the context of data re-purposing and reuse. Yet, data re-purposing represents an unprecedented opportunity for organisations to (re)create new value from existing data assets, highlighting the central importance of effective data sharing [39] and use [8] from the perspective of a socio-technical organisation.

Second, there is increasing evidence that data scientists spend over 80% of their time tackling problems related to data access, linkage, and cleaning [7]. There is a plethora of examples of situations where inadequate handling of data complexity has resulted in disastrous consequences. For example financial and reputational damage [15], and issues of social justice and public harm, such as propagation of biases in data into speech analyses [44], or data discrepancies and algorithmic assumptions that resulted in discrimination [17]. At the very least, the under-estimation of data curation and preparation needs results in time and cost over-runs in business analytics projects.

Third, data pipelines constructed by data engineers and scientists often lack explainability due to the use of complex computational, statistical, and machine learning techniques [34]. This causes, on the one hand, efficiency concerns due to poor transferable and repeatability prospects, and, on the other hand, deterioration in the trust of analytical results by business stakeholders and end-users. While literature on explainability [23] is advancing, a practice of transparency is still not that prevalent in the data science and analytics workforce [42]. These divides can create functional and cultural friction between the teams, and a disconnect between how consumers think their data should be handled and how it is actually treated [6]. While legislative and regulatory frameworks are playing catch up, currently most are untested for adequacy, and may simply be viewed as an increased compliance burden on organisations (especially SMEs) thereby limiting the uptake of data-driven solutions and innovations.

Challenges aside, there is no doubt that data-driven organisations that overcome the socio-technical barriers will emerge as winners in an increasingly competitive environment across all major sectors. The question for BPM research, vendor and practitioner communities is: what role can BPM play in helping organisations gain value from their data assets? In the section below, we highlight some of these opportunities.

2 Can BPM Help?

One central question for organisations that intend to pervasively use data for value creation relates to structuring of the analytics teams in ways that can transform the organisation into a data-driven entity [11]. However, this organisation-wide transformation is challenging in practice due to existence of functional silos and difficulties in creating proximity between analytics teams and business or domain groups [36]. The traditional function-based approach is where a central unit is established to serve multiple business units with their various analytics needs [16], and this approach generally leads to a strained relationship between analytics-oriented and business groups, because the former can rarely meet every demand of the latter [29].

Instead, data-driven work must be rooted in a pervasive enterprise-wide approach in which analytics is woven into the fabric of the organisation. Such an approach demands breaking down organisational (unit) boundaries to facilitate analytics teams and business groups to develop a common language to work collaboratively, and iteratively, and to integrate their knowledge into improved data-driven solutions, products and processes. This approach ultimately results in close ties between analytics teams and business groups, common language and motivation to interact, which brings about organisation wide change. To enable this proximity, organisations typically go beyond setting up centres of excellence that focus on enterprise analytics capabilities to build cross-disciplinary analytics teams that embed themselves within process-oriented groups.

BPM Opportunity: Process orientation has long advocated the need for organisation-wide thinking and breaking down of functional silos, and has been shown to be positively associated with firm performance [20]. *How can process orientation research help set up organisations for data-driven work?*

Despite the importance of effective use, many firms struggle to achieve it. Historically we know that, "...effective use is one of the greatest challenges for BI [Business Intelligence] systems. … Despite increasing investments in BI systems, many organisations are still unable to attain the desired success … due to underutilization and ineffective use" [4]. One theory that explains how organisations can use data and analytics systems more effectively is the Theory of Effective Use (TEU) [8]. TEU suggests that effective use involves three dimensions: (a) transparent interaction: seamlessly accessing the representations offered by a system, (b) representational fidelity: obtaining more accurate representations from the system, and (c) informed action: taking actions based on faithful representations.

We know that data can contain unexpected insights, and hence effective use of analytics systems can help organisations reap unexpected gains. However, the speed with which these insights can be effectively used is critical. Agility in value creation from data can be challenged with the way in which analytical insights are delivered, embedded and utilized within business. Organisational capability to support data-driven process design and improvement becomes essential. IT business literature advocates for a process perspective on value of IT resources and capabilities [25, 28]. Traditionally, analytics team outputs might be static reports or dashboards that exist separate from business processes and might not meet the specific decision requirements of managers and employees. However, data assets and capabilities generate business value when they are embedded within organisational processes. The challenge is how these operational and informational capabilities and experiences are blended into an integrated application or platform that promotes user engagement and empowers users with evidential decision making [40].

BPM Opportunity: Agile value creation from data requires new approaches for process design and improvement, in which data-driven insights can be embedded into processes in a timely and agile way that facilitate, extend and improve analytics-driven user experiences. *How can process design and adaptation research help facilitate agile use of data-driven insights?*

A notable process design choice related to how organisations integrate the insights with existing processes or develop new tools and processes is between (1) the augmentation of users' capabilities with algorithmic insights and recommendations to undertake evidential decision making and (2) the oversight mechanisms under which algorithmic insights and recommendations could be monitored and contested. The counter part is the design of fully automated decision-making, where algorithmic agents decide and act independently. There is growing evidence that this can create tensions. While algorithms can perform structured tasks and process massive datasets in real time, humans usually fare better with less structured tasks, especially ones that require creativity and interpretation [5]. Optimally, human-machine configurations should leverage both agents' strengths in a complementary manner. However, finding the right balance between automation and human involvement is not easy and practical guidelines are still emerging [30].

Whereas, process mining has proven a valuable approach in providing insights into the actual business process from organisational and case perspectives, the bulk of the advancements, understandably, are on the support of BPM [33], where process mining can enable evidence-based BPM [2], be used as a tool for Delta analysis and conformance testing [1] or to detect discrepancies, improve process, and provide better support (e.g. in the (re)design and diagnosis phases) for BPM life-cycle [26]. However, the rich body of knowledge on process mining tools and techniques can be translated into several novel domains including those that support value creation from data. For example, process mining can also reveal how people and/or procedures actual work [1] and provide understandable models that enables experts to understand the actual workflow and to detect specific user behaviours patterns [13]. The literature highlights the value of using process mining to understand human behaviors [33]. For example, model understanding has explored using process mining on eye tracking data (i.e., one of the physiological variables used as a technique to reflect the changes in cognition [27]) to find reading patterns in hybrid processes of DCR-HR [3], sensemaking behaviors in dual artefacts of business processes and rules [9], on domain and code understanding tasks from the developers' interactions [24], as well as on discovering data workers interaction behaviors and strategies in finding data quality issues data curation work [18].

BPM Opportunity: Process mining offers a rich set of methods and tools that can be used to understand human behaviour processes as well as process that have a mix of automated and human tasks. *How can process mining help achieve the right balance between automation and human involvement in data driven processes?*

Related to this opportunity is another that stems from the concept of reference models. Reference models are blueprints of best practice with the aim of reusability. They were popularised in the early 90s (see e.g. [35] and, since then, have been applied in a broad range of contexts [14]. The use of reference models has been associated with several benefits, including process improvement outcomes and risk reduction [32]. All data work, from data curation through to the development of AI models, is in itself a process. There is evidence that suggests that data curation processes are currently done in an ad-hoc manner [18, 19] and that process mining is helping to uncover a closer to optimal approach [18]. High profile failures in developing AI solutions, contrasted against notable successes [43], also suggest that there are best practices to be learned from.

BPM Opportunity: Process reference models offer a blueprint for best practice and enable organisations to improve their process performance. *How can process reference models be used to capture and share best practice approaches to value creation from data?*

3 Is the problem worth solving?

Evidence based decision making is not a new concept and has been the flagship approach for many sectors such as clinical research [12], policy reform [46] and financial markets [22]. However, we highlighted above the characteristics and challenges in evidential decision making with big data and complex black-box algorithms. In fact, advancements in machine learning (ML) and artificial intelligence (AI) are being valued at contributing up to US$15.7 trillion [31] to the global economy by 2030. AI is enabled by data [45] and the need for robust mechanisms for 'generating, sharing and using data in a way that is accessible, secure and trusted' is clear. Indeed, data gone wrong is acknowledged as the biggest risk factor for AI and other emerging technologies [21]. Unless organisations can see business value in data-driven work, the opportunity for responsible [37] and agile value creation from data will not materialise.

The authors of this paper posit that BPM research and practice holds a significant amount of knowledge capital that can be harnessed to contribute to the problem of value creation from data. We call the BPM community to assemble behind this exciting and interesting challenge of our times.

Acknowledgments. This work is supported by the ARC Industry Training Centre for Information Resilience – CIRES - IC200100022.

References

1. Aalst, W.M.P. van der: Business alignment: using process mining as a tool for Delta analysis and conformance testing. Requir. Eng. 10, 3, 198–211 (2005).
2. Aalst, W.M.P. van der: Process Mining: Overview and Opportunities. ACM Trans. Manag. Inf. Syst. 3, 2, 1–17 (2012).
3. Abbad Andaloussi, A. et al.: Exploring the Understandability of a Hybrid Process Design Artifact Based on DCR Graphs. In: Reinhartz-Berger, I. et al. (eds.) Enterprise, Business-Process and Information Systems Modeling. pp. 69–84 Springer International Publishing, Cham (2019).
4. Ain, N. et al.: Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. Decis. Support Syst. 125, 113113 (2019).
5. Asatiani, A. et al.: Implementation of automation as distributed cognition in knowledge work organizations: Six recommendations for managers. In: 40th International Conference on Information Systems, ICIS 2019. Association for Information Systems (2020).
6. Australian Competition and Consumer Commission: Digital platforms inquiry - final report, https://www.accc.gov.au/publications/digital-platforms-inquiry-final-report, last accessed 2021/05/19.
7. Belkin, R., Patil, D.: Everything We Wish We'd Known About Building Data Products | First Round Review, https://review.firstround.com/everything-we-wish-wed-known-about-building-data-products, last accessed 2021/05/19.
8. Burton-Jones, A., Grange, C.: From use to effective use: A representation theory perspective. Inf. Syst. Res. 24, 3, 632–658 (2012).
9. Chen, T. et al.: Sensemaking in Dual Artefact Tasks – The Case of Business Process Models and Business Rules. In: Dobbie, G. et al. (eds.) Conceptual Modeling. pp. 105–118 Springer International Publishing, Cham (2020).

10. Daugherty, P.R., Wilson, H.J.: Human + Machine: Reimagining Work in the Age of AI. Harvard Business Press (2018).
11. Davenport, T.H., Bean, R.: Big Companies Are Embracing Analytics, But Most Still Don't Have a Data-Driven Culture, https://hbr.org/2018/02/big-companies-are-embracing-analytics-but-most-still-dont-have-a-data-driven-culture, (2018).
12. Engebretsen, E. et al.: Unpacking the process of interpretation in evidence-based decision making. J. Eval. Clin. Pract. 21, 3, 529–531 (2015).
13. Fernández-Llatas, C. et al.: Process Mining for Individualized Behavior Modeling Using Wireless Tracking in Nursing Homes. Sensors. 13, 11, 15434–15451 (2013).
14. Fettke, P. et al.: Business process reference models: survey and classification. In: Business Process Management Workshops. pp. 469–483 Springer (2006).
15. Gillespie, N. et al.: Achieving Trustworthy AI: A Model for Trustworthy Artificial Intelligence. The University of Queensland and KPMG, Australia (2020).
16. Gregory, R.W. et al.: IT consumerization and the transformation of IT governance. 39.
17. Hajian, S. et al.: Algorithmic bias: From discrimination discovery to fairness-aware data mining. Presented at the ACM SIGKDD ICKDDM.
18. Han, L. et al.: On Understanding Data Worker Interaction Behaviors | Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.
19. Hey, T., Trefethen, A.: The Data Deluge: An e-Science Perspective. In: Berman, F. et al. (eds.) Wiley Series in Communications Networking & Distributed Systems. pp. 809–824 John Wiley & Sons, Ltd, Chichester, UK (2003).
20. Kohlbacher, M., Reijers, H.A.: The effects of process-oriented organizational design on firm performance. Bus. Process Manag. J. 19, 2, 245–262 (2013).
21. Korolov, M.: AI's biggest risk factor: Data gone wrong, https://www.idginsiderpro.com/article/3254693/ais-biggest-risk-factor-data-gone-wrong.html, last accessed 2021/05/19.
22. Leuz, C.: Evidence-based policymaking: promise, challenges and opportunities for accounting and financial markets research. Account. Bus. Res. 48, 5, 582–608 (2018).
23. Li, X.H. et al.: A Survey of Data-driven and Knowledge-aware eXplainable AI. IEEE Trans. Knowl. Data Eng. 1–1 (2020).
24. Loannou, C. et al.: Mining reading patterns from eye-tracking data: method and demonstration. Softw. Syst. Model. 19, 345–369 (2020).
25. Melville, N. et al.: Review: Information technology and organizational performance: An integrative model of IT business value. MIS Q. 28, 2, 283–322 (2004).
26. Netjes, M. et al.: Supporting the full BPM life-cycle using process mining and intelligent redesign. Contemp. Issues Database Des. Inf. Syst. Dev. 100–132 (2007).
27. Paas, F. et al.: Cognitive Load Measurement as a Means to Advance Cognitive Load Theory. Educ. Psychol. 38, 1, 63–71 (2003). https://doi.org/10.1207/S15326985EP3801_8.
28. Pang, M.-S. et al.: IT Resources, Organizational Capabilities, and Value Creation in Public-Sector Organizations: A Public-Value Management Perspective. J. Inf. Technol. 29, 3, 187–205 (2014).
29. Peppard, J.: Rethinking the concept of the IS organization. Inf. Syst. J. 28, 1, 76–103 (2018).
30. Rai, A. et al.: Editor's comments: next-generation digital platforms: toward human-AI hybrids. MIS Q. 43, 1, iii–x (2019).
31. Rao, A.S., Verweij, G.: Sizing the prize: what's the real value of AI for your business and how can you capitalise? PricewaterhouseCoopers Australia (2017).
32. Rosemann, M., van der Aalst, W.M.: A configurable reference modelling language. Inf. Syst. 32, 1, 1–23 (2007).
33. Rubin, V.A. et al.: Process mining can be applied to software too! In: Proceedings of the 8th ACM/IEEE International Symposium on Empirical Software Engineering and

Measurement - ESEM '14. pp. 1–8 ACM Press, Torino, Italy (2014).
34. Samek, W. et al.: Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. 1, 10 (2017).
35. Scheer, A.-W.: Business Process Engineering: Reference Models for Industrial Enterprises. Springer-Verlag, Berlin Heidelberg (1994).
36. Someh, I. et al.: Reconceptualizing synergy to explain the value of business analytics systems. J. Inf. Technol. 34, 4, 371–391 (2019).
37. Stoyanovich, J. et al.: Responsible data management. Proc. VLDB Endow. 13, 12, 3474–3488 (2020).
38. Surbakti, F.P.S. et al.: Factors influencing effective use of big data: A research framework. Inf. Manage. 57, 1, 103146 (2020).
39. Wixom, B.H. et al.: Data Sharing 2.0: New Data Sharing, New Value Creation, https://cisr.mit.edu/publication/2020_1001_DataSharing_WixomSebastianGregory, last accessed 2021/05/18.
40. Wixom, B.H., Ross, J.W.: How to Monetize Your Data. MIT Sloan Manag. Rev. 58, 3, (2017).
41. Zhang, R. et al.: Discovering Data Quality Problems. Bus. Inf. Syst. Eng. 61, 5, 575–593 (2019).
42. Beijing Consensus on Artificial Intelligence and Education - UNESCO Digital Library, https://unesdoc.unesco.org/ark:/48223/pf0000368303, last accessed 2021/05/19.
43. Delivering AI programs in the public sector: guidelines for government leaders, https://www.sap.com/documents/2020/11/48c79bd1-b97d-0010-87a3-c30de2ffd8ff.html, last accessed 2021/05/19.
44. TayTweets: Microsoft AI bot manipulated into being extreme racist upon release - ABC News, https://www.abc.net.au/news/2016-03-25/microsoft-created-ai-bot-becomes-racist/7276266, last accessed 2021/05/19.
45. The effective and ethical development of artificial intelligence: An opportunity to improve our wellbeing, https://acola.org/hs4-artificial-intelligence-australia/, last accessed 2021/05/19.
46. Tools for Institutional, Political, and Social Analysis of Policy Reform: A Sourcebook for Development Practitioners. World Bank Publications (2007).