BUSINESS PROCESSES MANAGEMENT AND RESPONSIBLE USE OF BIG DATA

Galena Pisoni[1] and Bálint Molnár[2]
[1]*Université Cote d'Azur*
Polytech Nice Sophia, Campus ShopiaTech, 930 Route des Colles, 06410 Biot, France
[2]*Eötvös Loránd University (ELTE)*
Information Systems Department, Faculty of Informatics Pázmány Péter 1/C, 1117 Budapest, Hungary

ABSTRACT

Organizations today seek to improve and adapt their business processes because of the competitive economy. The use and application of Data Science (DS) and Artificial Intelligence (AI) for different business process aims is often discussed and put in place, yet it can have a negative impact if DS and AI are implemented in the wrong way. We discuss the use of DS and AI for the identification, verification, and improvement of business processes, and how to ensure responsible DS and AI use in enterprises for this aim. We propose an information system design for responsible and trustworthy business processes, and we envision that businesses will need strong and well-defined control points in their information systems for managing processes and creating associated audits to enforce their principles. We define questions and challenges companies need to reflect on and follow for an application of responsible DS and AI in the enterprise context.

KEYWORDS

Business Processes, Business Process Management (BPM), Responsible AI, Data Science, AI, Enterprise Architecture

1. INTRODUCTION

The fast evolution of business process management systems and architectures has brought a generation of a waste amount of data coming from business process execution (Weske, 2007). These data can be used to discover or validate knowledge regarding existing processes operators have or help improve processes and provide flexibility and possibility for human intervention. A business process is "a collection of tasks executed in a specific order to achieve the business goal" (Weske, 2007). This definition suggests that business process management systems track the execution of the process it supports and ensures that the process confirms to set of restrictions. The application of data science (DS), and artificial intelligence (AI), opens the possibility to knowledge workers to obtain further knowledge, previously not known, and provides additional flexibility without endangering the correct execution of the business processes itself.

The use of DS and AI is on rise in companies. Data science (DS) studies how to collect, structure, and transform data into information and knowledge, it is an umbrella term for statistical techniques, design techniques, and development methods, while AI is the implementation of a predictive model to foresee events, it has to do with predictive algorithm design, development, efficiency, and the deployment in products and services (Waller & Fawcett, 2013; Newman et al., 2016). In business contexts, DS has termed business intelligence and analytics (BI&A), while the application of AI is referred to as artificial intelligence and analytics (AI&A). Both BI&A and AI&A support the collection, management, analysis, and visualization of large amounts of data, which is of great importance for providing managers, especially in terms of their own operations and business processes, with recommendations and insights (de Medeiros et al., 2020).

The topic of the application of DS and AI to business process management is relevant because organizations that use their business operations data to drive their decisions through BI&A and AI&A can achieve better performance (Brynjolfsson et al., 2011; Müller et al., 2018). According to Newman et al. (2016), the ability to manage operations with DS and AI has become significant for organizations, therefore also the adoption of the BI&A and AI&A approaches for this aim.

Organizations have understood the power of data analytics for mining and enhancing business processes, however, there are great concerns about the use of data for this aim as well. However, the power inherent in DS and AI naturally poses certain threats, especially in terms of fairness, trust, and privacy (Eitel-Porter, 2021; Van Der Aalst, 2016). DS and AI projects could deliver erroneous or biased outcomes and thus breach regulation or cause damage to the company reputation. These outcomes can happen due to different reasons, like bias in the data, bias encoded in the algorithms, or irresponsible management of both, data, and algorithms (Martin, 2022). Data decisions may be unfair or non-transparent, and confidential data may be shared unintentionally or abused by third parties (van der Aalst et al., 2017). Inaccuracies can be produced in each step of the DS or AI analysis, that is if the data used to train the model reflects biases, the output is likely to incorporate these biases as well (Van Der Aalst, 2016). Therefore, detection and prevention of bias and discrimination is an important aspect when applying DS or AI, especially for companies, where their brand and compliance with regulations at stake.

In the context where business processes and business process management are supported by DS and AI, all business processes should be able to run as intended and preserve the ability to recover and retract from undesirable and inconsistent outcomes. Interactive dashboards and business intelligence interfaces, in relation to the execution and adaptation of business processes, and options on how to mitigate problems observe, together with the suggestion of countermeasures, should help the knowledge operator with options and information regarding the business process at hand and provide new insights not available before.

To be able to exploit to the best of the potential the use of DS and AI, the business process management system should be able to interact with agents in an optimized manner. Implementing big data-enhanced business processes in companies and organizations whose information systems had not been originally structured for structuring and processing of big data, or that do not have management structures that enable big data analysis and value creation through the generation of actionable insights, is a big challenge. Moving to data-driven business process management requires both, suited personnel to apply DS and AI in the first place, a change in the underlying architectures and systems to allow for suited collection and analysis of big data, and a change in mindset.

To be able to interact with humans in the best possible way is an important characteristic of a business process management system enhanced with DS and AI. There are numerous ways in which DS and AI can help in this respect, they can provide more and new insights about the execution context and should be able to suggest process-related goals and actions, based on observed situational information, and should be structured so that can work and function in synchronous or asynchronous input. In addition, ideally, they can adapt the mode of working based on the preferences of the business user, adapt the information provided to the type of user it interacts with, and the situation in which the interaction happens. Acceptance of such DS and AI-empowered business process will depend on if the users find the advice rational and justified for the situation; thus, users should be able to understand the provided information, and the system provides faithful explanations for the choices. DS and AI can only be effective if people trust the results, therefore should not be viewed as a black box, accountability and comprehensibility are essential for transparency and therefore trustworthiness of the outcome. (van der Aalst et al., 2017).

Responsible data science implementations must ensure the fulfillment of the following principles (van der Aalst et al., 2017)

– Fairness, that is outputs should avoid unfair conclusions even if they are true
– Accuracy, that is the outputs and results should be accurate, and not a guesswork
– Confidentiality, that is results should answer questions without revealing secrets
– Transparency, that is outputs should provide an explanation to the final user on how the decision was made, and that explanation is clear and makes sense

For use of AI, there’s the AI ethics guideline for trustworthy AI by an EU Independent High-Level Expert Group on Artificial Intelligence, that highlights the need for AI systems to be human-centric. The new EU regulatory framework applies to products and services relying on AI. Ethics Guidelines for Trustworthy AI advocate for AI technology that is more human-centric and doesn’t discriminate, therefore also responsible and trustworthy business processes based on big data should follow these guidelines and behave in compliance with them.

Improper implementations of AI and DS can bring systematic discrimination and invasions of privacy, therefore it’s important to prevent non-transparent and inaccurate conclusions, and to avoid the negative side effects that data may have in any context (van der Aalst et al., 2017).

Given all the above, in this article we first hypothesize an information system design for responsible and trustworthy business processes, next we outline, still to answer challenges, research questions, and opportunities for research in the domain, and finally, we outline actors and stakeholders that should work together to answer the outstanding challenges and questions in this domain. We conclude the paper with a short discussion/reflection of outstanding questions and outline future directions.

2. BACKGROUND

In order to have a responsible and trustworthy application of DS and AI in a business processes context, there must be adequate tools, methods, and technologies to support such governance and decision-making based on big data (Chen et al., 2012). These elements, that is governance frameworks and technologies, should come together, so responsible DS and AI principles are applied at each stage of big data analysis of business operations context.

Given the new demands of companies to constantly innovate and to stay relevant, the research in the domain of information systems is united that IS are considered as 'never complete' and constantly need to be improved (Tilson et al., 2010; Tiwana et al., 2010). The development of governance of business processes, based on big data, requires that sociotechnical constructions are revisited, and new decision points, based on data and observations from outputs of the application of DS and AI methods are introduced. These new strategic decision control points, in return, introduce new stable socio-technical constructions. The new strategic decision points require re-development of the management platforms, and there's a process of flexibility/stability that follows until a new IS stable configuration is found (Figure 1).

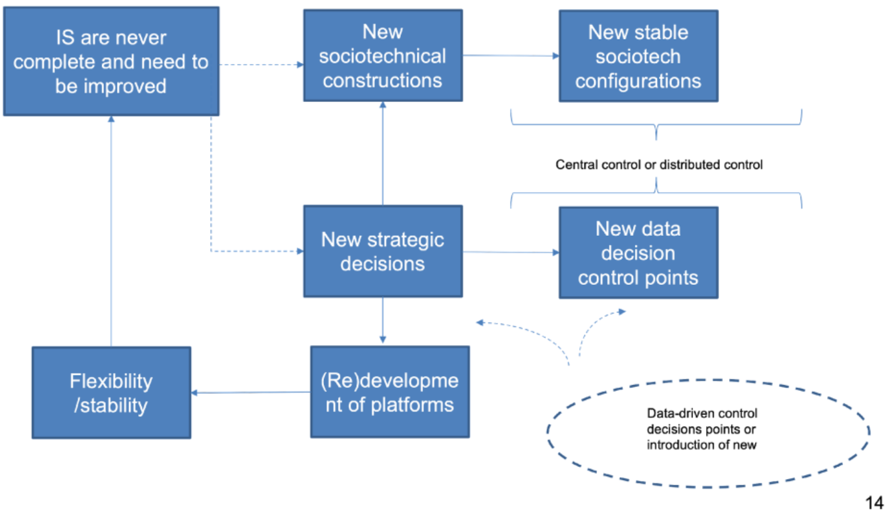

Figure 1. Process of governance of business processes management, with new sociotechnical constructions and decision points decided on big data and application of AI and DS, inspired by Tilson et al. (2010)

The new IS design to support such business process execution, requires that IS architecture is modular by design, decomposed, and follows simple design rules, the design, in turn, must have a good fit with IS use and governance for business operations, therefore, decision points and control mechanisms, are decided based on data, and application of DS and AI on operations data. This IS organization is influenced by the environment in which the company operates and the technological trajectory of the products and services of the adjacent companies and competitors, as well as their influence (Figure 2). The business issues request data warehouse or data lake design suited to the requirements, therefore they also drive competitors in a similar way in terms of IS design and search for optimal solutions for the domain at hand.

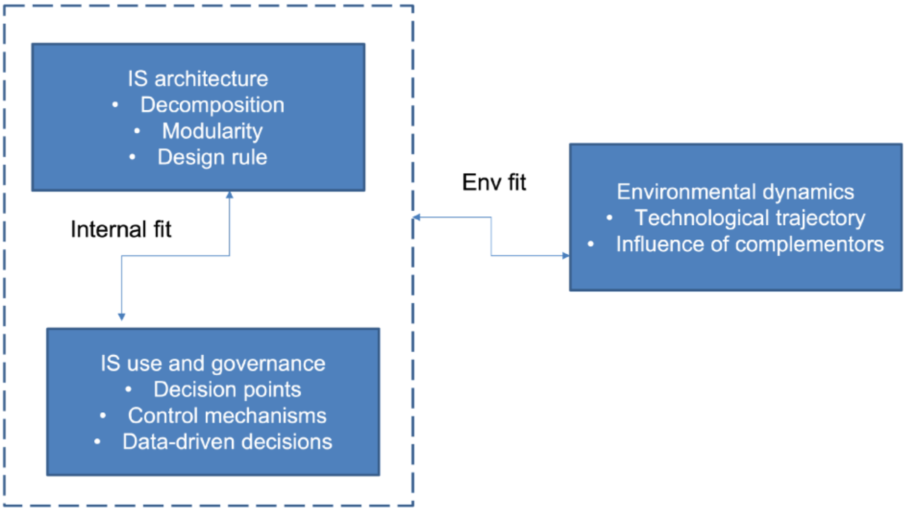

Figure 2. IS architecture is influenced internally by the use and governance of IS, and new decision points, based on data, while the whole IS organization and management is decided in the wider environmental context, in which the competitors and adjusts technologies affect the different selections internally of the company for its operations, inspired by Tiwana et al. (2010)

Ethical DS and AI governance in businesses foresee that first the company determines, based on existing guidelines, the ethical principles that apply to its own work and to which it will adhere. Preferably there should be an ethical board, that among other important members of the company, consists of ethicists, and approve the strategies then implemented in DS and AI implementations, especially for own company operations. As proposed in our framework, there must be a clear outlines governance structure, based on data collected and the application of AI and DS, there must be clarity over who is expected to take decisions about which aspects of the business operations, and possible tradeoffs and benefits from the company, as deduced from the AI and DS analysis. Next, such knowledge should be spread across the whole company and all the company members, and finally, there should be strong monitoring, preferably with metrics, if the outputs of both, data-driven approaches and governance do give ethical, responsible, and trustworthy results

3. IMPLEMENTING ETHICAL AND RESPONSIBLE BUSINESS PROCESSES

3.1 Guidelines for Companies

The problem as we outline in this paper lies in the intersection of human-centric AI and enterprise architectures and applications, with a focus on business process optimization for businesses. In the previous section, we outlined a governance model for IS for business process management, however, different questions and challenges need to be answered before such a governance model, enhanced with AI and DS, is put in place in organizations.

Explainable AI (XAI), for instance, promises to address the explainability problem, that is to enable users to understand the algorithm and parameters it used to make the decision. The aim of XAI is to provide the reasons for the outcome of the process of deciding a way to make it understandable for the final user. The research on XAI had mainly been performed in two directions: visualization of ML processes and explainable ML algorithms. These approaches so far have delivered limited results, due to the algorithms' limited capability to explain decisions in humanly understandable ways or rely on abstract visualization methods, which may further increase complexity for final users.

In an organization, there are different kinds of users and stakeholders, and different stakeholders may need different explanations. In addition, not everyone or every role needs or wants an explanation, it may be that the same explanation may be good for more than one role (Langer, M et al., 2021). Therefore, the first question that companies will need to answer in this respect is

1) Which stakeholders need an explanation? What kind of explanation each stakeholder/user needs?

Subsequently, there'll be the need to test the effectiveness of the explanations generated and how users understand the explanations. There's a need for participatory involvement of all and co-creation of meanings for different users, that is developers of DS and AI solutions and business experts, the need to also teach business users regarding explainability and how to "consume" explanations generated, risks involved, and common ways in which the explanations will be used (Chromik et al, 2021). After being provided an explanation, some participants may want or need even more information, on how one can do this efficiently the companies will need to understand of their own:

2) How can we test if the explanation is good for the needs of the organization? How can we provide more information for the roles/stakeholders that need such additional information? How can we achieve better alignment of the explanations with the needs of the different stakeholders in the company?

A known problem in DS is the phenomenon of "overfitting", or better looking for patterns that don't generalize outside of the given dataset, in addition, AI algorithms fail as well, and there are risks concerned with the actions stakeholders usually take or might need to take, so another question is

3) How risky AI or DS failures are for the company?

If users understand how DS and AI work, then they'll understand also preventively what kind of problems may arise, for companies. It may be easier to provide access to people with such knowledge that can later be explained, especially when the explanatory information is not good or is superficial, not useful; moreover, they are not available, as sometimes to make the results understandable, when the results are good, can be achieved after consultation of such people.

Therefore, the other two questions to consider for companies:

4) Do different stakeholders understand the explanations produced by AI or DS?

5) Is it possible to have DS and AI scientists available on spot to provide explanations to different stakeholders?

Sometimes explanations provided are not the desired one and may be inadequate, for deciding in the complex world of business. Sometimes the explanation can be of no-good quality, or too detailed, in different formats, and different kinds of contents, like diagrams, matrices, and logic trees (Wang et al, 2019). Previous research confirms that it's good to visualize data or make a reduction but clear analogies, especially in problem-solving domains, in highly detailed technical explanations are not favored often, although having cross-disciplinary skills would help the users to feel more secure in the results as they will be able to control the outputs and the data better in their own.

6) What should we do if the provided explanations are too technical?

In the use of DS and AI approaches and data in general it may be better to also know better the data origin, and the data itself and to see the –input-output demonstration, and therefore understand the biases

7) How will we provide more information about the data origin? In which format?

In a context like this, it is important also for all stakeholders to understand how the system can fail or mislead.

8) Can we identify a common way in which the system can fail? Can we provide such training to different stakeholders in the company?

Interactive ML is an approach to ML in which algorithms interact with agents (these agents can be humans too) and thus through these interactions they can optimize their behavior (Holzingher, 2016) for a problem like this, the use of interactive ML (iML) is a key as well. It may be the case that the traditional ML approaches might have limited success, in the case of events not available in the training dataset.

The next challenge for the company to solve would be which data should be used for developing such a system. There are privacy-and security-related consequences, as well as, as outlined above ethical issues that are relevant and important to be considered (Clavell & Peuvrelle, 2020). There is the observatory of algorithms with social impact studies that enlists different algorithms and provides information about their aim, therefore aiming to demystify opaque and unaccountable algorithms that have proved to be systematically biased against women and minority groups, this can be another valuable resource regarding selection of algorithms on which the system can be based.[1]

In addition, companies need to be compliant with regulations, therefore the system needs to comply with regulations too, especially guidelines coming from the EU regarding the development of human-centric AI systems (European Commission, 2019). There are also existing widgets to use for the development of a fair and not biased decision-making system (based on e.g., fairness IBM Toolbox and responsible AI widgets) that can make the development of AI solutions easier for companies.

The methods for the analysis of data should be mixed and encourage ethical and socially responsible behavior. Computational social science approaches and computational analytics methods based on machine learning and computational models to contribute to comprehensible AI, taking into account the business processes and business needs should be designed, to accelerate the transition to actionable, responsible, and bias-free insights from big data. There is a need of further research to evaluate the practices regarding knowledge management and data analysis, as well as the gathering, and usage of data in the generic business process management ecosystem of the company, to test the application of relevant existing tools, widgets, and DS and AI methods so to have good quality output.

3.2 What Kind of Team to Gather?

The presented challenges and questions for companies require knowledge and expertise from different domains, and form a team combining a diverse set of needs to answer them:

– Data custodians, data curators, and knowledge operators, that would prepare the data coming from different operations for analysis, especially from big data and data coming from different sources of the company

– Data scientists, that would use the generated data from business process execution to build robust machine learning models that are responsible, trustworthy, and explainable

– Business managers, that are familiar with the knowledge domain of the company, and existing business processes the company relies upon,

– Regulation authorities, to confirm that the explanations on how the decision was made, are in line with regulation if it is clear and makes sense

– Ethicists, and company external references regarding ethical criteria in the domain of the company.

4. DISCUSSION AND FUTURE DIRECTIONS

In this paper, we outlined governance mechanisms for business process management, based on big data, DS and AI, and a set of guidelines for companies, in the form of a checklist to ensure good implementation of DS and AI-enhanced business processes. In addition, we reflect on what kind of team to gather to be sure that there is ethical and responsible use of AI and DS for business process improvement.

There are many possible future directions and problems still to solve.

One of the issues in research like this is which decisions should be fully automated (based on data) and for which decisions human input shall be needed and why? And how can companies decide on this? An important step would be to understand and distinguish functions that should be carried out by DS or AI-empowered systems, and functions that should be performed entirely by humans, and study their order in each context, to understand how intelligent functions can best support human actions, and vice versa.

[1] https://eticasfoundation.org/oasi/register/

Another challenge is how to put together the same visualization insights generated from data and the process itself. Another recurring challenge is the relationship between data and processes. State-of-the-art approaches like UML still have different diagrams covering different aspects, that is, class models for data, activity diagrams for processes, yet there are no suited visualizations that combine data, business intelligence insights, and business process representations.

Petri nets fully integrate both but are difficult to use for complex designs of information systems. Further research in the domain is timely.

ACKNOWLEDGEMENT

The paper was supported by the “Application Domain-Specific Highly Reliable IT Solutions" project, implemented with the support provided by the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme TKP2021-NVA-29 (National Challenges Subprogramme) funding scheme, and by the COST Action CA19130 - "Fintech and Artificial Intelligence in Finance Towards a transparent financial industry" (FinAI) https://www.cost.eu/actions/CA19130/.

REFERENCES

Brynjolfsson, E., Hitt, L. M. & Kim, H. H. (2011), ‘Strength in numbers: How does data-driven decision-making affect firm performance?’, Available at SSRN 1819486 .

Chen, H., Chiang, R. H. & Storey, V. C. (2012), ‘Business intelligence and analytics: From big data to big impact’, MIS quarterly pp. 1165–1188.

Chromik, M., Eiband, M., Buchner, F., Krüger, A., & Butz, A. (2021, April). I think i get your point, AI! the illusion of explanatory depth in explainable AI. In *26th International Conference on Intelligent User Interfaces* (pp. 307-317).

Clavell, G. G., & Peuvrelle, V. (2020). Ethical Issues in Big Data Analytics for Time Critical Mobility Forecasting. In Big Data Analytics for Time-Critical Mobility Forecasting

de Medeiros, M. M., Hoppen, N. & Maçada, A. C. G. (2020), ‘Data science for business: benefits, challenges and opportunities’, The Bottom Line .

Eitel-Porter, R. (2021), ‘Beyond the promise: implementing ethical ai’, AI and Ethics 1(1), 73–80.

European Commission (2019). High-level expert group on Artificial Intelligence, Ethics Guidelines for Trustworthy AI, European Commission

Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human- in-the-loop?. Brain Informatics, 3(2), 119-131

Langer, M et al. (2021). What do we want from Explainable Artificial Intelligence (XAI)?–A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research. *Artificial Intelligence*, 296, 103473

Martin, K. (2022), Algorithmic bias and corporate responsibility: How companies hide behind the false veil of the technological imperative, in ‘Ethics of data and analytics’, Auerbach Publications, pp. 36–50.

Müller, O., Fay, M. & Vom Brocke, J. (2018), ‘The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics’, Journal of Management Information Systems 35(2), 488–509.

Newman, R., Chang, V., Walters, R. J. & Wills, G. B. (2016), ‘Model and experimental development for business data science’, International Journal of Information Management 36(4), 607–617.

Tilson, D., Lyytinen, K. & Sørensen, C. (2010), ‘Research commentary—digital infrastructures: The missing is research agenda’, Information systems research 21(4), 748–759.

Tiwana, A., Konsynski, B. & Bush, A. A. (2010), ‘Research commen- tary—platform evolution: Coevolution of platform architecture, governance, and environmental dynamics’, Information systems research 21(4), 675–687.

Van Der Aalst, W. (2016), Process mining: data science in action, Vol. 2, Springer.

van der Aalst, W. M., Bichler, M. & Heinzl, A. (2017), ‘Responsible data science’. Waller, M. A. & Fawcett, S. E. (2013), ‘Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management’.

Weske, M. (2007), Business process management architectures, Springer.

Wang, D., Yang, Q., Abdul, A., & Lim, B. Y. (2019, May). Designing theory-driven user-centric explainable AI. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-15).