

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

“ Listening to the data is important... but so is experience and intuition. After all, what is intuition at its best but large amounts of data of all kinds filtered through a human brain rather than a math model? ”

Steve Lohr
Technology and economics journalist

Analyzing “big data” holds huge potential for generating business value. The ongoing advancement of tools and technology over recent years has created a new ecosystem full of opportunities for data-driven innovation. However, as the amount of available data rises to new heights, so too does complexity. Organizations are challenged to create the right contexts, by shaping interfaces and processes, and by asking the right questions to guide the data analysis. Lifting the innovation potential requires teaming and focus to efficiently assign available resources to the most promising initiatives. With reference to the innovation process, this article will concentrate on establishing a process for analytics projects from first ideas to realization (in most cases: a running application). The question we tackle is: what can the practical discourse on big data and analytics learn from innovation management? The insights presented in this article are built on our practical experiences in working with various clients. We will classify analytics projects as well as discuss common innovation barriers along this process.

Introduction

Understandably, much effort is being expended into analyzing “big data” to unleash its potentially enormous business value (McAfee et al., 2012; Wamba et al., 2017). New data sources evolve, and new techniques for storing and analyzing large data sets are enabling many new applications, but the exact business value of any one big data application is often unclear. From a practical viewpoint, organizations still struggle to use data meaningfully or they lack the right competencies. Different types of analytics problems arise in an organizational context, depending on whether the starting point is a precise request from a department that only lacks required skills or capabilities (e.g., machine learning) or rather it stems from a principal interest in working with big data (e.g., no own infrastructure, no methodical experience). So far, clear strategies and process for value generation from data are often missing.

Much literature addresses the technical and methodical implementation, the transformative strength of big data (Wamba et al., 2015), the enhancement of firm performance by building analytics capability (Akter et al., 2016; Wamba et al., 2015), or other managerial issues (Davenport & Harris, 2007; McAfee et al., 2012). Little work covers the transformation process from first ideas to ready analytics applications or in building analytics competence. This article seeks to address this gap.

Analytics initiatives have several unique features. First, they require an explorative approach – the analysis does not start with specific requirements as in other projects but rather with an idea or data set. To assess the contribution, ideation techniques and rapid prototyping are applied. This exploration plays a key role in developing a shared understanding and giving a big data initiative a strategic direction. Second, analytics projects in their early phase are bound to a complex interplay between

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

different stakeholder interests, competencies, and viewpoints. Learning is an integral part of these projects to build experience and competence with analytics. Third, analytics projects run in parallel to the existing information technology (IT) infrastructure and deliver short scripts or strategic insights, which are then installed in larger IT projects. Due to a missing end-to-end target, data is not only to be extracted, transformed, and loaded, but also needs to be identified, classified, and partly structured. So, a general process for value generation needs to be established to guide analytics projects and address these issues.

Here, we propose an exact configuration and series of steps to guide a big data analytics project. The lack of specified requirements and defined project goals in a big data analytics project (compared to a classic analytics project) make it challenging to structure the analytics process. Therefore, the linear innovation process serves as reference and orientation (Cooper, 1990). As Braganza and colleagues (2017) describe, for big data to be successfully integrated and implemented in an organization, clear and repeatable processes are required. Nevertheless, each analytics initiative is different and the process needs to be flexible. Unfortunately, the literature rarely combines challenges in the analytics process with concepts from innovation management. Nevertheless, an integration of the concepts from innovation management could guide the analytics work of formulating digital strategies, organizational anchoring of the analytics units and their functions, designing the analytics portfolio, as well as the underlying working principles (e.g., rapid prototyping, ideation techniques).

Thus, in this article, we will concentrate on the question of what the practical discourse and work on analytics respectively implementing big data in organizations

can learn from innovation management. A process for analytics innovation is introduced to guide the process from ideation to value generation. Emphasis is put on challenges during this process as well as different entry points. Thereby, we build on experience and insights from a number of analytics projects for different sectors and domains to derive recommendations for successfully implementing analytics solutions.

We begin with a definition of big data and analytics. Next, we propose a process for a structured approach to retrieving value from data. Finally, we discuss the results and outline directions for future research.

Big Data and Analytics

In this section, we address the elementary angles from which the analytics value chain should be looked at (Figure 1): data, infrastructure, and analytics – and the business need as the driver. According to our understanding, value is generated by analyzing data within a certain context, with a problem statement related to a business requirement driving the need for innovation. Besides expertise in conducting data and analytics projects, this process requires a working infrastructure, especially when volume, velocity, or variety of data to be analyzed exceeds certain limits. Below, we describe the three technical angles in more detail.

Data

Big data is often defined with volume (how much data), velocity (speed of data generation), and variety as the diversity of data types (Chen & Zhang, 2014; Gandomi & Haider, 2015). Big data describes data collections of a size difficult to process with traditional data management techniques. While many definitions of big data concentrate on the aspect of volume referring to the

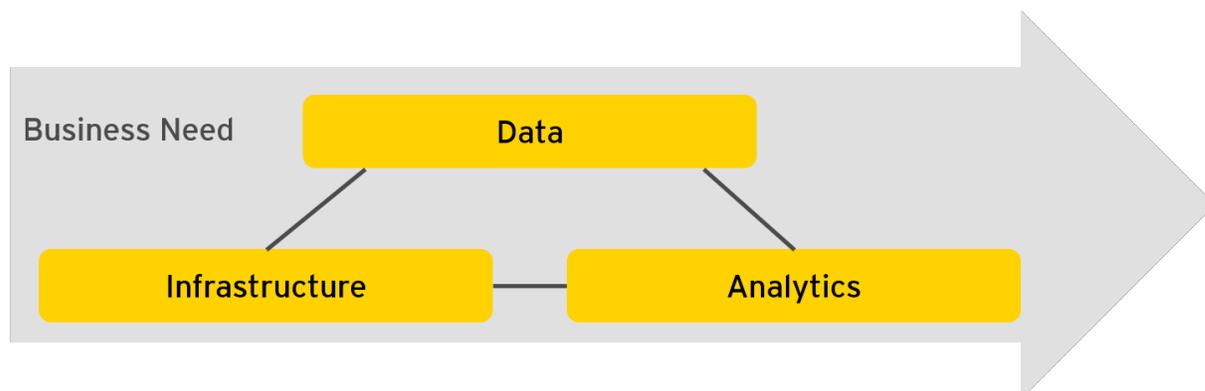


Figure 1. Framework of data, infrastructure, analytics and business need

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

scale of data available, big data brings in particular heterogeneous formats and a broad spectrum of possible data sources. Examples are structured numeric data or unstructured data such as text, images, or videos. This variety and broad landscape of data sources offers many opportunities for generating insights. Moreover, the speed of data creation enables rapid insights in ongoing developments.

Recent technical improvements (e.g., cloud computing, big data architectures) enable data to be analyzed and stored on a large scale. For many (new) types of data, their exact business value is unclear so far and requires systematic exploration. Available data is often messy, and even when cleaned up can be overwhelming and too complex to be easily understood, even by professional data scientists. The contribution of data is, of course, context specific and varies among business cases and applications. One key challenge is to identify data that best meets the business requirement.

Analytics

Data science is concerned with knowledge generation from data. Analytics or data science addresses the exploration of data sets with different quantitative methods motivated from statistical modelling (James et al., 2015) or machine learning (Mitchell, 1997). Methods from different disciplines such as statistics, economics, or computer science find application to identify patterns, influence factors, or dependencies. In contrast to business intelligence, analytics reaches further than descriptive analytics (based on SQL) and often has a predictive component. Which method to apply depends on the exact business case. Analyzing data is restricted, for example, by a company's internal policies as well as legal restrictions and guidelines that vary among countries. Data quality and reliability are further issues. Data understanding and domain knowledge are key prerequisites in the analysis process (e.g., Waller & Fawcett, 2013), especially when model assumptions are made.

Concerning data analysis, there are primarily the following opportunities for organizations:

- Improved analysis of internal data: One example is forecasting methods that enhance expert-based planning approaches by additional figures. These methods build on existing databases such as business intelligence systems, and they contribute new or further insights to internal firm processes.
- New combinations of data sets offer new insights, for example, through the combination of sensor data and user profiles.
- Opening up to new or (so far) unused data sources (e.g., websites, open data) to identify potential for generating new insights. However, a context or application is necessary to use the data. One example is social media data used for market observation.

However, the core problem of analytics is to work out the guiding question and achieve a match between business need, data source, and analysis as discussed later in the article.

IT infrastructure

Relevant for the successful implementation of analytics is the adaption of the IT infrastructure to embed analytics solutions and integrate different data sources. The core layers of an IT infrastructure are the following:

1. **Data ingestion layer:** This layer covers the data transfer from a source system to an analytics environment. Therefore, a toolset and a corresponding process need to be defined. Traditional extract, transform, load (ETL) tools and relational databases are combined with Hadoop/big data setups covering, in particular, scenarios caused by less structured, high volume, or streamed data. Analytics use cases build on data from data warehouses to fully unstructured data. This breadth challenges classic architectures and requires adaptable schemes. Which data sources to integrate depends on the specific application.
2. **Data value exploration layer:** Based on the business need and corresponding use case, data is investigated, tested, and sampled in this layer. Depending on the complexity and business question, an appropriate analytics scheme is developed. Business and explorative analysis based on online analytical processing (OLAP) models in memory technologies are supplemented or expanded by using advanced analytics methods and integrating (e.g., R or Python plugins).
3. **Data consumption layer:** Here, the results are used for visualization, for example. The end user can consume the data or service without deep technical understanding (e.g., for self-service business intelligence).

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

Modern approaches require structures that are adaptable and scalable to different requirements and data sources. Factors such as system performance, cost efficiency, and overall enterprise infrastructure strategy must be taken into consideration.

From Data to Value: Turning Ideas into Applications

Organizations still struggle to use data meaningfully or lack the right competencies. One of the key challenges in analytics projects is identifying the business need and the guiding questions. Principally, different types of analytics problems arise in an organizational context ranging from precise requests that only lack specific capabilities to a principal interest in working with big data (e.g., no own infrastructure, expert-based approaches). This approach implies different starting points for the analytics process and different innovation pathways, both of which are described later in this article.

What is the starting point?

The starting point for each analytics initiative varies. According to the four points mentioned above, the “state of the art” for each one needs to be assessed individually to estimate the analytics maturity:

- 1. Business need:** From case to case, the precision of the problem description and scope varies. For some cases, the leading question and scope guiding the analysis phase are formulated very precisely and for other cases it needs to be worked out and refined during the process.

- 2. Data:** The data to be used in the project can be defined or an appropriate source is not yet clear. The size and quality of the data essentially determine the progress of the further process. Parameters are, for example, structure (i.e., pre-processing effort) or the size of the data set (e.g., one CSV file or a large database).
- 3. Analytics:** Which methods to apply differs from case to case and must be tested and explored.
- 4. Infrastructure:** The current (technical) state of the business unit (e.g., own data warehouse, reporting system) or own (human) resources and competencies is a further important aspect in classifying requests.

These four angles can be rated differently with reference to the maturity level of the analytics request. Based on our experience, three scenarios, representing different maturity levels, can be distinguished (Figure 2):

- In scenario 1, the data analysis is motivated by a defined requirement such as market observation during the rollout of a new product. The appropriate data source needs to be identified. The data missing so far implies that the precise analysis cannot be defined and also that there is no existing infrastructure. Ideas need to be developed as to which data sources could be relevant and which issues can be resolved on this basis. Then, different methods from data analysis are applied to generate new insights.
- In scenario 2, the data source and infrastructure are clearly defined, and the specific questions need to be

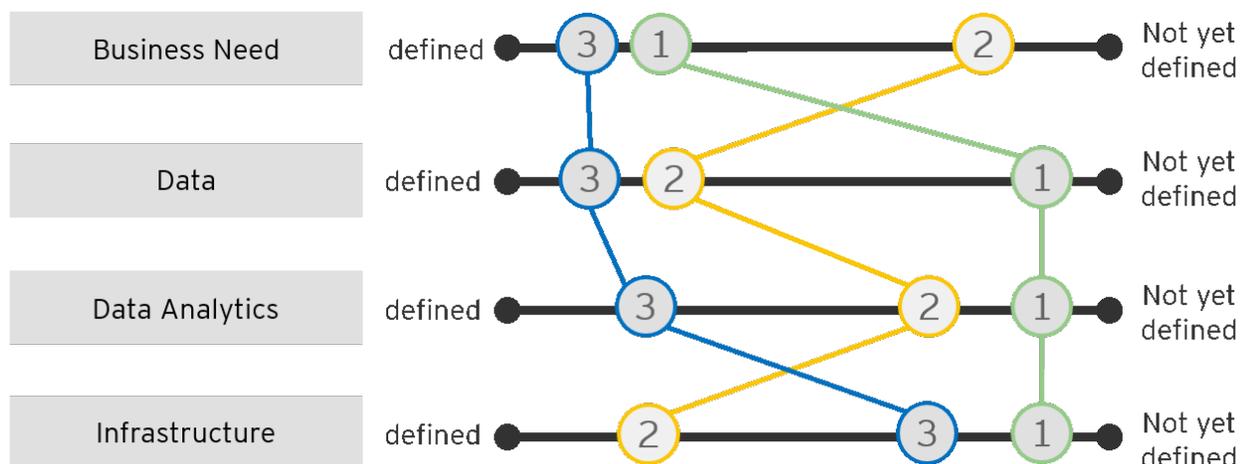


Figure 2. Classifying analytics requests: Three maturity levels

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

identified. One application is assessing the contribution of a specific data source that has not been professionally analyzed so far, for example, by means of machine learning. For instance, the business unit has an internal database, considers new methods, and wants to further develop a business intelligence system by adding a forecasting component. In this case, the scope is clearer than in the first scenario and straight away an explorative data analysis can be started.

- In scenario 3, there is a precise analytical problem that needs to be professionalized. A first draft shows promising results and the solution can, as a next step, be upscaled. Guidance in making architectural decisions is needed.

These three scenarios are exemplary starting points for analytics projects. The following section describes the implications for the innovation process and outlines different challenges and barriers.

The analytics process

To succeed with analytics, the process from data to value must be structured to be integrated in the existing organization. For example, Braganza and colleagues (2017) examine the management of organizational resources in big data initiatives. They stress the importance of systematic approaches and processes to operationalize big data.

Related work on analytics processes has a focus on service design (Meierhofer & Meier, 2017) or concentrates on the methodical part of analyzing data (e.g., Cielen & Meysman, 2016). The process, as introduced by Braganza and colleagues (2017), is too linear and does not address the systemic complexity of data analysis and necessary stakeholder discourse. To cover these issues, structuring the analytics process can be linked to the classic linear innovation process (Cooper, 1990; Salerno et al., 2015).

In our work, to guide the analytics process from ideation, scoping, and identifying a data set to value generation, a process with four phases is introduced. Taking the classic innovation funnel as starting point, this concept is transferred to the context of analytics. The process is divided in four parts: i) idea generation, ii) proof of concepts (PoCs) are conducted to test these ideas, iii) the successful PoCs are implemented and tested, and, finally, iv) they become available as a product or service. Based on a first idea or requirement, the process is initialized, while the number of ideas or projects is reduced within each phase. Each phase has tasks as well as barriers or filters that need to be passed to continue in the process chain.

The three scenarios described above are assessed differently concerning their maturity as illustrated in the process in Figure 3. Scenario 1 is in a very early stage of

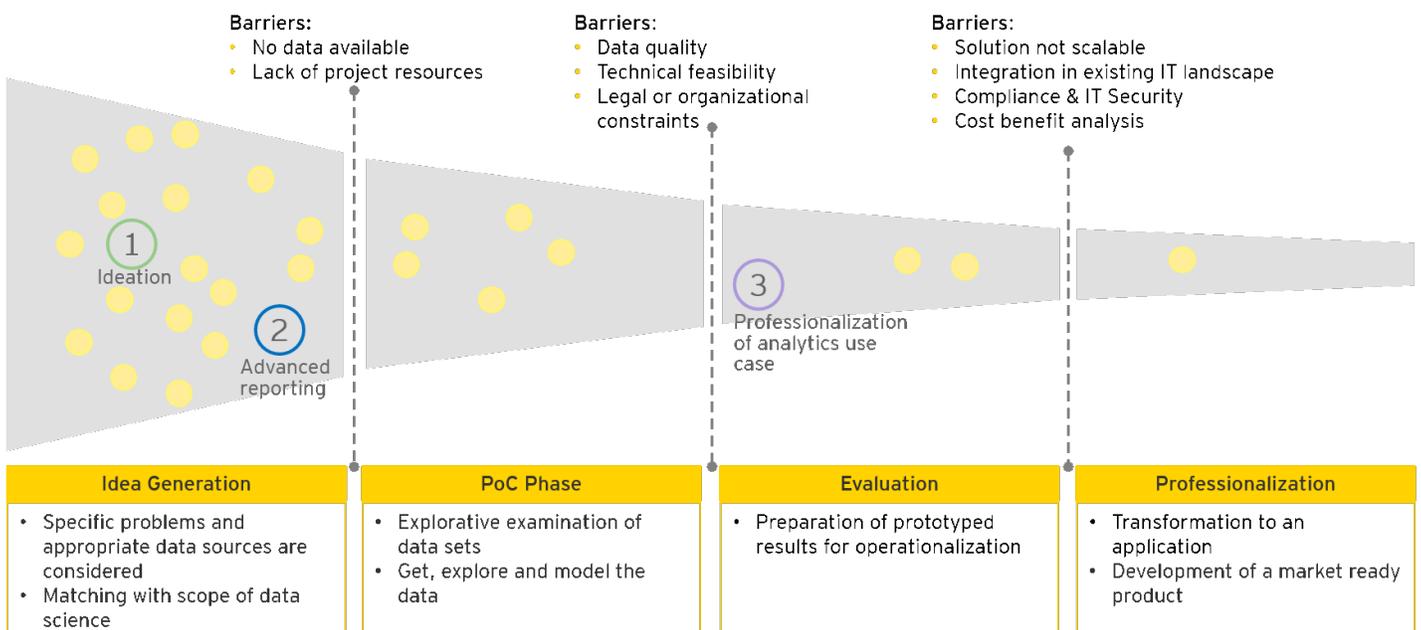


Figure 3. Phases of the analytics process

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

idea generation and many open questions need to be addressed. Scenario 2 is more concrete and many more issues are resolved than in scenario 1. However, initiating questions need to be developed before a PoC can be conducted. Scenario 3 builds on a running system, so it is located in the phase of testing and operationalization (phase 3).

For each phase, different challenges arise. While related work emphasizes data-related challenges such as data acquisition, cleansing or aggregation (Sivarajah et al., 2017), this work focuses on process challenges.

Phase 1: Idea generation

Orientating analytics projects begins with an ideation phase. Here, the key challenge is to gather ideas and discuss relevant business problems (see also Provost & Fawcett, 2013). Idea generation plays a key role in developing a shared understanding, challenging existing assumptions, orientating big data initiatives, and identifying aspects that can be solved with analytics. For example, design thinking is applied as a systematic approach to problem solving (Liedtka & Ogilvie, 2011) and supports a structured ideation process. Problems of the business unit are collected and matched with the scope of analytics (e.g., technical feasibility, input parameters, and methodical requirements). The ideation phase is iterative. Initially, the general project object-

ives guide the first ideation round, which aims at getting an overview of present challenges and needs of the business unit. This is in line with identifying appropriate data sets. Then, the feasibility of the ideas must be checked by experts and the ideas are then selected for prototyping.

From an organizational perspective, involvement of decision makers from all hierarchy levels is a must. Top management is required to resolve conflicts of interest and to create a sense of urgency, middle management is required to free experts from daily work and onboard stakeholders into their particular roles, whereas the expert knowledge of operative specialists is key to detailing the guiding question and checking the feasibility.

A portfolio is drawn to select the ideas that are considered in the PoC phase. Innovation portfolios provide a coherent basis for judging the possible impact of ideas (Tidd & Bessant, 2013). They separate ideas into areas and indicate which ideas to prioritize. For the exemplary case as illustrated in Figure 4, the ideas are rated and assessed according to three categories: feasibility (x-axis), value creation (y-axis), and overall relevance (size of the node). *Feasibility* contains aspects such as data availability, time to access data, or the expected complexity of the task. *Value creation* addresses the expected business value and underlines ideas with

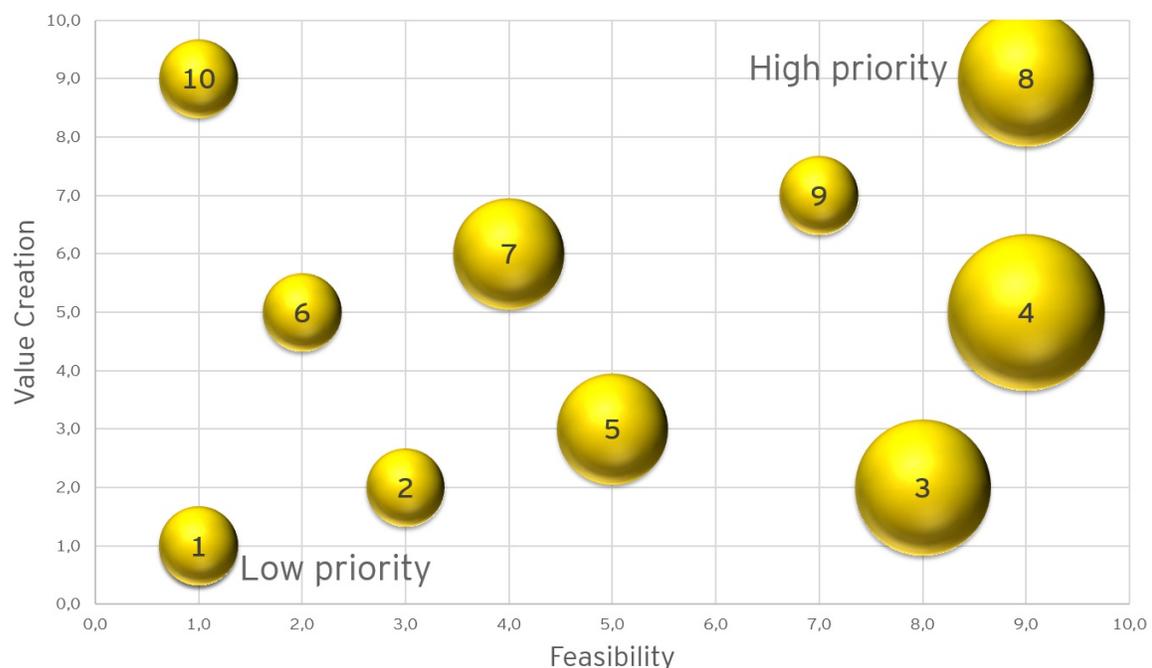


Figure 4. Portfolio for selecting ideas

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

a high expected contribution. The *overall relevance* is used to emphasize which ideas are expected to have greater impact on the problem at hand. So, for example, idea 3 has a high expected feasibility, but the created value is expected to be low. By contrast, idea 4 and idea 8 are bound to a higher expectation concerning value creation and should therefore be prioritized in the next phase.

Besides the portfolio-based selection process, ideas are filtered during the first phase, for example, because there is no data available to address the problem, the data must be raised first (e.g., implementation of additional sensors), or access is denied (e.g., internal policies, legal restrictions). So, appropriate data sources need to be identified and access needs to be granted for a reliable yet efficient assessment of business needs and data applicability.

As an organizational barrier, the right experts need to be identified and freed of their daily work such that they are available for analytics projects. During the ideation process, the right balance between creativity and focus is important as well as bridging the gaps between diverse knowledge areas to ask the right questions.

The outcome of this first phase are ideas plus the data sources on which basis the problems can be examined; a mapping of problems or ideas and data sources is required. In the first phase, strong facilitators are needed

to guide through the process. In addition, someone with methodical expertise to check the technical feasibility of the ideas considered as well as business understanding are important. The ideas and data are only discussed; no examination takes place. This is done in the next step.

Another issue that needs to be clarified in this early phase are data security and data protection. Each country has individual regulations that limit the analysis.

Phase 2: Proof of concept

To test the ideas, prototypes are built and PoCs are conducted. PoCs are a first examination of the data set to see if a raised question can be answered based on the available data or not.

This phase is described in Figure 5: based on the defined scope from the previous phase, access to the data must be granted, the data is explored and analyzed, and finally the results are communicated.

As described above, this phase begins with a project goal or problem description (business need). Whereas classic IT development starts with requirements, analytics often starts in an explorative way with a dataset and a hypothesis. Specific requirements are generated during the analysis process. So, the PoC phase can only start with data or when data is available. Getting the data or retrieving it from existing systems is among the first steps in a PoC. Here, access barriers such as legal issues or organizational constraints need to be checked.

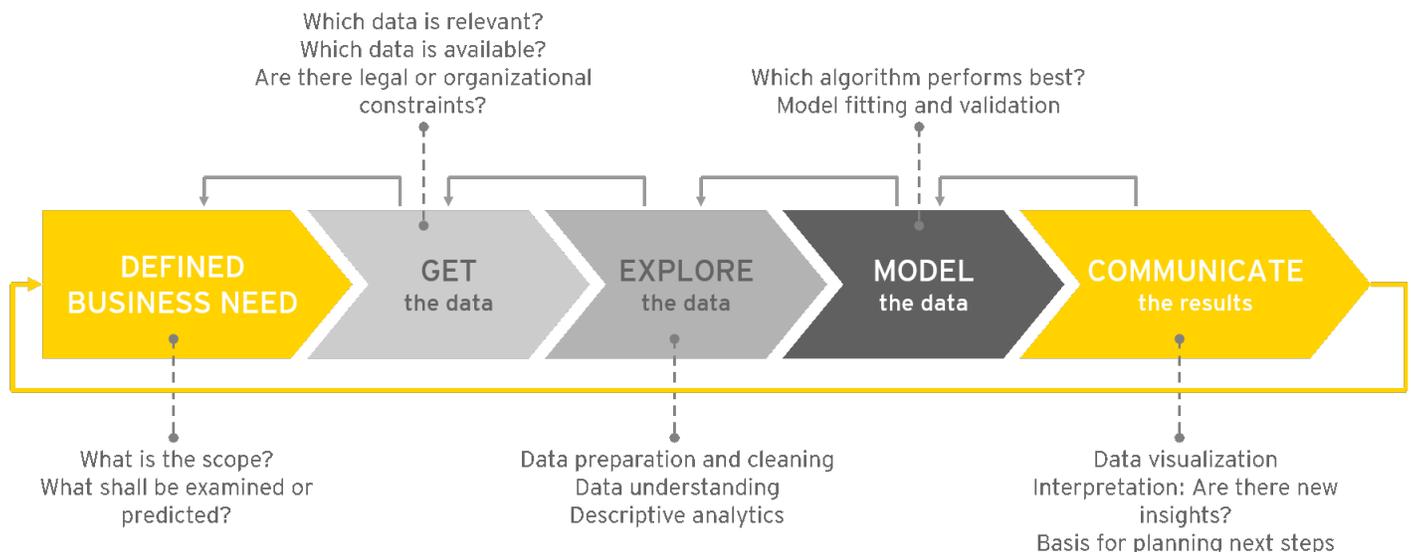


Figure 5. The analytics process

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

So, for example, depending on the type of data (e.g., personal data, machine data, market figures), the analysis should be in line with these restrictions.

Next, the data is explored for a deeper understanding. Here, the data is transformed to a suitable format for further analysis. This step contains data preparation and cleaning, and the first descriptive analysis is conducted.

The data is then analyzed for patterns and dependencies during the modelling phase to answer the questions raised. Different methods and algorithms are tested and the results are validated in an iterative process of variable selection, model selection, model adaptation, and validation.

Finally, the results are communicated. A PoC gives a first orientation on the potential in the data with an emphasis on strengths and weaknesses. Possible results are that different modelling techniques do not deliver a valid result, the data quality does not allow modelling, or there is not enough data for a significant statement. This is finally the basis for planning and communicating next steps and coordinating further actions.

Concerning the presentation of the results, different visualization techniques can be applied working with tools like such as Tableau, QlikView or different open source platforms. Especially to develop an understanding of the data, descriptive data analysis is helpful. Nevertheless, many models and techniques from advanced analytics deliver figures that cannot be captured by intuitive visualizations.

PoCs have a short duration of maybe 6–8 weeks. Besides getting access to the necessary data or extracting data from relevant sources, among the key challenges in this phase are data quality, data ownership, and data understanding. Further barriers are cleansing and munging of the data to a format that can be processed and to apply the right models. Furthermore, business understanding is key to retrieving valuable insights from the data and achieving outcomes that are not only plausible but relevant for the business. Another issue is the lack of experience with analytics and the required agility in implementing the results.

Phase 3 & 4: Operationalization

Then, the PoC results are integrated into a professional IT infrastructure. Prototyped results need to be prepared for operationalization and transformed to an ap-

plication. The main question to answer is: Is the model scalable and can results achieved so far be applied to a larger data set? Adjustments have to be made so that a resulting application can be maintained by an IT service organization without continued support from data scientists. Event or time-based data flows have to be established and, together with the final application, need to be aligned with compliance, security, and data privacy requirements. Test management and service-level agreements for incident handling and application changes need to be agreed on as well as product and portfolio management functions in case the tool or application is meant to assume a strategic, long-term role.

Barriers include, for example, the required budget, overly complex tests, standards, and compliance. Together, the integration in IT management and allocation of tasks to the IT department represent another issue. This relates to switching from an agile, iterative working model to stable operations and scaling the analytical model and transferring it to maintainable code.

Generally, great effort is required in transforming the PoC prototype into a professional infrastructure. Further barriers during operationalization are, for example, establishing support and service management functions, achieving acceptance among the user base, developing adequate training concepts, and transferring knowledge required to maintain, test, and develop the application.

Discussion and Conclusion

Generally, the challenge for organizations lies in defining strategies for value generation from the large amount of available data sets. In this article, we discussed how to retrieve value from data and introduced a systematic process that analytics projects follow. First, we described the fundamental building blocks for value creation: business need, data, infrastructure, and analytics. Then, we described the process from ideation to market ready applications. According to the maturity state of the project, the process can be entered at different stages. The four phases of this process were described with emphasis on the specific barriers. This model is oriented towards a stage-gate model (Cooper, 1990) for analytics processes and aims to structure and systematize explorative analytics approaches.

Analytics and big data are not only a technical challenge but impact the whole organization and its processes. For being successful with analytics, less effort

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

should be spent on building complex and sophisticated models but instead on integrating the results into the existing (technical) infrastructure and processes. For the prototype being professionalized, the results must be accepted and understood, and the business unit should be continuously involved in the process. Moreover, the right set of people and skills is necessary: not only are data scientists with competencies in machine learning and statistical modelling required (Mikalef et al., 2017), but also IT specialists and business understanding in general. In addition, value is only generated from data if the analysis is integrated into an overall framework of skills and competencies and the analytics initiative is embedded in a business application.

The results of this article can be transferred to organizations of different sizes and levels of experience when building analytics capabilities. The process as described in this work guides through analytics projects and illustrates the differences to known IT management approaches. By principally discussing the meaning of innovation for analytics, this work contributes to the evolving literature on digital innovation management (Nambisan et al., 2017). In our work, we have outlined an approach for data-driven innovation.

Future work should examine the decisions in organizing analytics. This covers aspects as roles and responsibilities, team structures, leading analytics teams, or the organizational embedding of analytics units in the organization. The results of this work should be linked to the extensive research on analytics capability, which are often classified along the dimensions of management, technology, and human capability (Akter et al., 2016; Mikalef et al., 2017). Throughout the process, as introduced in this work, the understanding of analytics becomes clearer. So, its contribution to organizational learning, skill development, developing a shared understanding, and building analytics capability should be examined. For example, according to Davenport and Harris (2007), this analytics learning process needs around 18–36 months. From a technical point of view, in particular the integration of analytics solution into the overall IT landscape, the professionalization of prototypes and change of established processes remain challenging.

Acknowledgements

This article was developed from a paper presented at the ISPIM Innovation Conference in Vienna, Austria, June 18–21, 2017. ISPIM (ispim-innovation.com) – the International Society for Professional Innovation Management – is a network of researchers, industrialists, consultants, and public bodies who share an interest in innovation management.

About the Authors

Victoria Kayser is a Data Scientist in Ernst & Young's Advisory Organization. Her research is focused on the intersection of analytics and innovation management. Her PhD examined the contribution of text mining to foresight and future planning. She has worked in the fields of innovation research and strategy development as well as in the automotive sector. She holds a Master of Science degree in Information Engineering and Management.

Bastian Nehrke is a Manager with Ernst and Young's Advisory in Stuttgart. He specializes in developing organizational analytics capabilities and supports clients in setting up their own analytics hubs and CoEs as well as innovation and data thinking methods. He is a certified Project Manager, Business Analyst, and Requirements Engineer and studied International Management and Innovation and Technology Management in Frankfurt and Heilbronn.

Damir Zubovic leads Ernst and Young's Data and Analytics Practices in Germany, Switzerland, and Austria as Partner. With 15 years of professional experience in leading business intelligence, analytics and big data initiatives, he is responsible for business development, specializing in analytics applications in the automotive and life sciences sectors and in consumer and retail products. His extensive experience in the field makes him an experienced mentor who also acts as coach, lecturer, and keynote speaker.

Data Science as an Innovation Challenge: From Big Data to Value Proposition

Victoria Kayser, Bastian Nehrke, and Damir Zubovic

References

- Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. 2016. How to Improve Firm Performance Using Big Data Analytics Capability and Business Strategy Alignment? *International Journal of Production Economics*, 182: 113–131. <https://doi.org/10.1016/j.ijpe.2016.08.018>
- Braganza, A., Brooks, L., Nepelski, D., Ali, M., & Moro, R. 2017. Resource Management in Big Data Initiatives: Processes and Dynamic Capabilities. *Journal of Business Research*, 70: 328–337. <https://doi.org/10.1016/j.jbusres.2016.08.006>
- Chen, C. L. P., & Zhang, C.-Y. 2014. Data-Intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data. *Information Sciences*, 275: 314–347. <https://doi.org/10.1016/j.ins.2014.01.015>
- Cielen, D., & Meysman, A. D. B. 2016. *Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools*. Shelter Island, NY: Manning Publications.
- Cooper, R. G. 1990. Stage-Gate Systems: A New Tool for Managing New Products. *Business Horizons*, 33(3): 44–54. [https://doi.org/10.1016/0007-6813\(90\)90040-1](https://doi.org/10.1016/0007-6813(90)90040-1)
- Davenport, T. H., & Harris, J. G. 2007. *Competing on Analytics: The New Science of Winning*. Boston, MA: Harvard Business School.
- Gandomi, A., & Haider, M. 2015. Beyond the Hype: Big Data Concepts, Methods, and Analytics. *International Journal of Information Management*, 35(2): 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. 2015. *An Introduction to Statistical Learning: With Applications in R* (6th ed.). New York: Springer.
- Liedtka, J., & Ogilvie, T. 2011. *Designing for Growth: A Design Thinking Tool Kit for Managers*. New York: Columbia Business School Publishing.
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. 2012. Big Data: The Management Revolution. *Harvard Business Review*, 90(10): 60–68.
- Meierhofer, J., & Meier, K. 2017. From Data Science to Value Creation. In *Proceedings of the 7th International Conference on Exploring Services Science (IESS)*: 173–181. Bucharest, Romania, May 25–27, 2016.
- Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. 2017. Big Data Analytics Capabilities: A Systematic Literature Review and Research Agenda. *Information Systems and e-Business Management*, 17(2): 1–32. <http://dx.doi.org/10.1007/s10257-017-0362-y>
- Mitchell, T. M. 1997. *Machine Learning*. New York: McGraw-Hill.
- Nambisan, S., Lyytinen, K., Majchrzak, A., & Song, M. 2017. Digital Innovation Management: Reinventing Innovation Management Research in a Digital World. *MIS Quarterly*, 41(1): 223–238.
- Provost, F., & Fawcett, T. 2013. *Data Science for Business*. Sebastopol, CA: O'Reilly.
- Salerno, M. S., de Vasconcelos Gomes, L. A., da Silva, D. O., Bagno, R. B., & Freitas, S. L. T. U. 2015. Innovation Processes: Which Process for Which Project? *Technovation*, 35: 59–70. <https://doi.org/10.1016/j.technovation.2014.07.012>
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. 2017. Critical Analysis of Big Data Challenges and Analytical Methods. *Journal of Business Research*, 70: 263–286. <https://doi.org/10.1016/j.jbusres.2016.08.001>
- Tidd, J., & Bessant, J. 2013. *Managing Innovation: Integrating Technological, Market and Organizational Change* (5th ed.). New York: John Wiley & Sons.
- Waller, M. A., & Fawcett, S. E. 2013. Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2): 77–84. <http://doi.org/10.1111/jbl.12010>
- Wamba, S. F., Akter, S., Edwards, A., Chopin, G., & Gnanzou, D. 2015. How 'Big Data' Can Make Big Impact: Findings from a Systematic Review and a Longitudinal Case Study. *International Journal of Production Economics*, 165: 234–246. <https://doi.org/10.1016/j.ijpe.2014.12.031>
- Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J.-f., Dubey, R., & Childe, S. J. 2017. Big Data Analytics and Firm Performance: Effects of Dynamic Capabilities. *Journal of Business Research*, 70: 356–365. <https://doi.org/10.1016/j.jbusres.2016.08.009>

Citation: Kayser, V., Nehrke, B., & Zubovic, D. 2018. Data Science as an Innovation Challenge: From Big Data to Value Proposition. *Technology Innovation Management Review*, 8(3): 16–25. <http://doi.org/10.22215/timreview/1143>



Keywords: big data, idea generation, innovation process, analytics, digital innovation

Academic Affiliations and Funding Acknowledgements



The Federal Economic Development Agency for Southern Ontario (FedDev Ontario; feddevontario.gc.ca) is part of the Innovation, Science and Economic Development portfolio and one of six regional development agencies, each of which helps to address key economic challenges by providing regionally-tailored programs, services, knowledge and expertise.

- *The TIM Review receives partial funding from FedDev Ontario's Investing in Regional Diversification initiative.*



Carleton
UNIVERSITY



Technology Innovation Management (TIM; timprogram.ca) is an international master's level program at Carleton University in Ottawa, Canada. It leads to a Master of Applied Science (M.A.Sc.) degree, a Master of Engineering (M.Eng.) degree, or a Master of Entrepreneurship (M.Ent.) degree. The objective of this program is to train aspiring entrepreneurs on creating wealth at the early stages of company or opportunity lifecycles.

- *The TIM Review is published in association with and receives partial funding from the TIM program.*