A Framework for the Systematic Evaluation of Data and Analytics Use Cases at an Early Stage

Dustin Schwarz
DAIN Studios
dustin.schwarz@dainstudios.com

Michael List
Bundesdruckerei
michael.list@bdr.de

Roland M. Mueller
Berlin School of Economics and Law
roland.mueller@hwr-berlin.de

Abstract

Due to the immense growth of collected data and advancing big data technologies, there are countless potential use cases of data and analytics. But most data initiatives fail and do not bring the desired outcome. One essential reason for this situation is the lack of a systematic approach to evaluate and select promising analytics use cases. This study presents an evaluation framework that enables the systematic screening at an early stage by assessing nine criteria with the help of a scoring model. It also supports a prioritization among several use cases and facilitates the communication to decision makers. The action design research approach was followed to build, test, and evaluate the framework in three iterative design cycles. It was developed in close collaboration with Bundesdruckerei GmbH, an IT-security company owned by the German government that offers products and services for secure identities, data, and infrastructures.

Keywords: Analytics use case evaluation, business value of analytics, data idea screening, analytics evaluation framework

1. Introduction

Today, we live in the age of analytics where organizations are competing in a data-driven world. The improvement of computational speed and storage, the exponential growth of available data and the development of more sophisticated algorithms are fueling this trend as new technologies advance. If used in the right way, those technologies allow companies to accelerate and outperform competitors regardless of their size. This could even lead to the disruption of entire industries as business models are changing or completely new ones evolve. For this reason, almost every organization explicitly mentions data as a critical enterprise asset and analytics as an essential competency (White, 2019). But not even half of the organizations are managing their data as an asset and even less operate in a data-driven way (NewVantage Partners, 2021).

Most of the organizations fail to benefit from their investments in data initiatives (Colas et al., 2014, p. 3; VentureBeat, 2019; White, 2019). The majority of the use cases are actually never developed and just a few are considered as successful by executives. The reasons for this situation are diverse. One essential factor for success is a systematic and structured approach to evaluate and select a suitable use case based on well-defined criteria. But many organizations are lacking such an evaluation process (Colas et al., 2014, p. 7). This situation can lead to errors in decision making, where either an unbeneficial use case is funded, or a great opportunity is missed. This usually results in high costs and a waste of resources (Baker & Albaum, 1986, p. 32).

This paper addresses the gaps by answering the following research question: *How can data and analytics use cases be evaluated at an early stage, with the help of an evaluation framework?*

The research is conducted in collaboration with the manufacturing department and data experts of Bundesdruckerei GmbH. Among other things, the department is responsible for the production of ID documents, driving licenses, banknotes, and postage stamps.

The framework facilitates an early-stage evaluation by including criteria that allow a credible assumption about the potential and feasibility of a data and analytics use case. In addition, the framework should be understandable and easy to use, even if the user's data literacy level is low.

2. Related work

New technologies and tools that are able to deal with big data support data-driven innovation. The term *data-driven* indicates that the processing and analyzing of (big) data is part of the innovation process (Kayser et al., 2018, p. 16).

Traditionally, innovation projects follow four phases: idea generation, idea selection, development, and implementation (Hansen & Birkinshaw, 2007; Salerno et al., 2015, p. 3). The first two phases are often referred to as the *fuzzy front end* which has a major

URI: https://hdl.handle.net/10125/103292
978-0-9981331-6-4

impact on the whole innovation process (J. Kim & Wilemon, 2002; Moon & Han, 2016, p. 84).

Kayser et al. (2018) present an innovation process that addresses the problem of achieving a match between business need, data source, and analysis. Additionally, common barriers that can eliminate ideas in the different phases are included. A similar approach is introduced by Vanauer et al. (2015). They distinguish between two different perspectives on data-driven initiatives: *data-first* and *business-first*. Whereas the business perspective focuses on use cases for operational improvement, the data perspective aims at designing new services that could be sold to other entities.

The idea screening itself is a process that is based on defined evaluation criteria to select and prioritize use cases (Karger, 1983). The *stage-gate-system* introduced by Cooper (1990) even divides the screening into several stages that are equal to quality control checkpoints (gates) that an idea has to pass after each phase of the innovation process. Besides the number of stages, also the criteria and methods used for each stage can vary (Cooper, 2011; Rochford, 1991; Schmidt et al., 2009, p. 20).

However, the process should minimize errors in decision-making as those can lead to high costs (Baker & Albaum, 1986, p. 32; Hansen & Birkinshaw, 2007, p. 5). Those errors can be defined as situations in which decision-makers either miss an opportunity with high value (*Type I Error*) or fund a use case that leads nowhere (*Type II Error*) (Hammedi et al., 2011, p. 665; Rochford, 1991, p. 294). Furthermore, the process should be effective and efficient by enabling a rapid consensus and decision-making (Hammedi et al., 2011, p. 665). Gerlach & Brem (2017, p. 148) present an overview of factors that have a positive influence on the outcome like transparency and clear criteria.

One thing to choose is the evaluation method. While some are intuitive and based on gut-feeling, other concepts are more structured and formal. With enough domain expertise and clear rules, intuition can be a sufficient and less resource dependent way of screening ideas (Magnusson et al., 2014, p. 323). On the other hand, intuitive screenings may lead to the investment of unfeasible ideas (Tauqeer & Bang, 2019, p. 2). For this reason, a formal analysis is applied in many cases. Baker & Albaum (1986) sort the formal approaches into eight categories. But even those do not fully eliminate the subjectivity of intuitive approaches, as ratings and screening decisions are always highly uncertain and subjective (Cooper, 1981, p. 59).

Besides the used methods and number of screenings, use case evaluations can also differ in terms of the people that participate in the process. Individual reviews followed by a team review is a common way to do the screening (Karger, 1983, p. 46). Other evaluators could be senior managers (Cooper, 1990, p. 46), consumers (Toubia & Florès, 2007), or expert committees (Lauto & Valentin, 2016).

There is a big number of studies that are dealing with the question which evaluation criteria to use. They should be well selected as too tight screening criteria can shut down any innovation and too loose ones can result in the funding of ideas that are not fitting to the corporate strategy and lead nowhere (Hansen & Birkinshaw, 2007, p. 5). Many studies name similar criteria. An overview is presented in Table 1. According to Tauqeer & Bang (2019, p. 5), all potential screening parameters can be grouped into the following: *producibility, problem size, market size, novelty, business alignment, profit margin,* and *others*.

Table 1. Published evaluation criteria

Study	Criteria
Karger (1983, p. 46)	Expected market impacts, predicted financial results, required resources and capabilities
Cooper & Brentani (1984, pp. 153–154)	<u>Dominant:</u> financial potential, corporate synergy, technological synergy, product differential advantage <u>Secondary:</u> product life, size of market, diversification strategy, market maintenance strategy, domestic market
Baker & Albaum (1986, p. 35)	Societal factors, business risk factors, demand analysis, market acceptance factor, competitive factors
Cooper (1990, p. 52)	Strategic alignment, project feasibility, magnitude of the opportunity, differential advantage, synergy with the firm's core business and resources, market attractiveness
Rochford (1991, pp. 294–295)	Market, product, feasibility, fit to organization and management, time, financials, others (do-ability, probability of success)
Carbonell-Foulquié et al. (2004, p. 312)	Technical feasibility, strategic fit, customer acceptance, financial performance, market opportunity
Dean et al. (2006, p. 663)	Novelty, workability/feasibility, relevance, specificity
Sandström & Björk (2010)	Novelty, usefulness, risk, benefit, effort
Kudrowitz & Wallace (2013, p. 137)	Novelty, usefulness, feasibility
Magnusson et al. (2014, p. 323)	Originality, user value, producibility, strategic fit, profitability
Stevanovic et al. (2015, p. 7)	Technical factors, customer factors, market factors, financial factors, social factors
Yarmohammadi et al. (2017, p. 672)	Technical feasibility (technical knowledge, reasonable costs, open APIs), legal feasibility, market potential (application input, sales potential)

Sometimes, criteria are further classified into must and want objectives (Rochford, 1991, p. 292). The study of Yarmohammadi et al. (2017) specifies those general criteria and includes software relevant aspects like open APIs, or application input. Several studies present tools to support the evaluation process. There are simpler ones like the ladder diagram developed by Tauqueer & Bang (2019) or the matrices presented by Kim & Mauborgne (2000) that are used to assess the utility, price, and business model of an idea. Those are suitable for a quick and initial screening of a bigger number of ideas. In contrast to that, Huang et al. (2020) combined qualitative and quantitative methods to choose the best eco-innovative ideas out of many. Also, IT supported solutions were developed like the workflow designed by Ciriello et al. (2016).

Besides the frameworks that support the idea screening for a general innovation project, there are a few with a focus on data-driven innovation. These are usually collaborative and canvas-based tools like the *Data Innovation Board* (Kronsbein & Mueller, 2019) or the *Data Insight Generator* (Kühne & Böhmann, 2020). Those support the generation of data and analytics use cases. Other canvas-based frameworks like the *Data Science Canvas* (Neifer et al., 2021) or the *AI Project Canvas* (Zawadzki, 2020) focus on the detailed description of the use case itself. A framework that is not canvas-based, is provided by UNDP & UN Global Pulse (2016). They designed a guide including several tools to support the process of bringing an initial idea to a proof-of-concept.

Two major gaps are pointed out in the literature review. Even though various studies suggest methods and criteria to guide the evaluation process, they do not consider the obstacles that are related to data and analytics. Evaluation criteria are too unspecific when it comes to data and analytics use cases. On the other hand, tools that include the data perspective usually aim at the generation of new use cases. However, they neither show if those ideas are beneficial and feasible nor enable a comparison between them. A systematic use case evaluation is missing. For these reasons, none of the mentioned tools is suitable for evaluating data and analytics use cases.

3. Methodology

To create the evaluation framework, the *action design research* (ADR) methodology was followed (Sein et al., 2011). This method is used to create an IT artifact that is shaped in iterative design cycles by the organizational context (Hevner et al., 2004, p. 77; Sein et al., 2011, p. 37). By that, the artifact aims at solving organizational problems “that emerge from the interaction of people, organizations, and technology” (Sein et al., 2011, p. 37). The ADR method consists out of the following four steps: (1) problem formulation, (2) building, intervention, and evaluation (3) reflection and learning, as well as (4) formalization of learning (Sein et al., 2011, p. 41).

The addressed problem is the absence of a systematic approach to evaluate and prioritize data and analytics use cases. The building, intervention, and evaluation step was conducted in three design cycles including mainly qualitative interviews and a quantitative questionnaire for the evaluation.

The sessions were conducted in German. Relevant statements were immediately transcribed and translated into English afterwards. In the end, all translated statements were clustered and coded to facilitate the comparison and analysis. Table 2 provides an overview of the design cycles. The participants that took part in the sessions work for Bundesdruckerei GmbH. A purposeful sample of employees with different professional expertise and organizational roles was chosen to allow various perspectives on the framework. By that, the quality of the artifact evaluation is increased. To align the sample with the purpose of evaluating data and analytics use cases, only individuals with a certain level of data literacy were selected.

Table 2. Design cycle overview.

Cycle	Partic-ipants	Roles	Purpose
1	4	Data Scientist, Data Analytics Consultant, Business Developer, Industrial Engineer	Assessment of internal evaluation approaches
1	7	Data Analysts, Project Manager (Manufacturing)	Validation of evaluation criteria
2	1	Manufacturing Engineer	Testing and evaluation
3	4	Process owner, Business Developer, Project Manager	Demonstration and evaluation

The first design cycle started with qualitative interviews. The goal was to (1) assess the existing practices within the organization to evaluate data and analytics use cases and (2) to understand organizational requirements regarding the evaluation tool. Based on the interviews and literature research, potential evaluation criteria were derived. Additionally, a survey was conducted with potential users of the tool to check the significance of the criteria. The insights were used to build the initial framework.

Afterwards, two iterations (cycle 2 and 3) were conducted. Following the ADR principles, the feedback and learnings of both cycles were incorporated into the artifact to shape the design of the framework.

The first iteration included the testing of the framework on a data analytics use case of Bundesdruckerei GmbH. The use case aimed at using actuator data from the passport production process to predict the quality of the product. In case of a predicted defect, adjustments could be made in advance to avoid it. To evaluate this use case, the framework was filled out by the initiator. To discuss ambiguities and answer questions, the testing was supervised by one of the authors. Based on that, the initial design was adjusted, and the alpha version was developed.

In a second design iteration, the adapted alpha version was evaluated by other experts who are also potential users of the framework. As the generation of ideas is outside the scope of the paper, the participants were not able to test the framework on a self-developed use case. Therefore, the use of the framework was demonstrated by presenting the alpha version and the testing results from the previous iteration phase. The feedback was incorporated into the artifact to develop the final beta version.

Furthermore, a quantitative questionnaire was filled at the end of each iteration by the corresponding participants to assess the following evaluation criteria: *usability, usefulness, practicality, efficiency, and design*. The questionnaire to evaluate the framework includes mostly single-choice questions and Likert-scales to assess the beforementioned criteria. The criteria of usability, usefulness and practicality are based on Kayser et al. (2019, p. 9). As the evaluation tool should facilitate a quick consensus and decision, the criterion of efficiency was added. According to Hevner et al. (2004, p. 86), the artifact should also be aesthetically pleasing to the user. For this reason, the criterion of design was included.

4. Use case evaluation framework

The evaluation process suggested by Samset & Christensen (2017, p. 6) was used as guidance. At first, meaningful evaluation criteria were derived. Afterwards, those criteria were decomposed into more detailed evaluation questions that should be answered. With the help of the framework, the information is acquired to answer the corresponding evaluation questions. Finally, a way to aggregate the results was derived. The final output of the evaluation is used to draw conclusions and make recommendations regarding the data and analytics use case.

4.1 Selection of criteria

As shown in section 2, the literature is lacking evaluation criteria that focus on data and analytics use cases. Therefore, new criteria had to be derived. As a starting point, common process models of data and analytics use cases like the *Cross-Industry Standard Process for Data Mining* (CRISP-DM) (Chapman et al., 2000) or the more agile *Team Data Science Process* (TDSP) (Tab & Sharkey, 2020) were analyzed. Those process models emphasize the importance of the *business understanding* on the one hand and the *data understanding* on the other hand. Therefore, criteria from both perspectives need to be considered when evaluating data and analytics use cases.

As explained in Section 2, there are some criteria that are mentioned in several studies like novelty, market potential, strategic fit, and feasibility. Those were used as a base to derive criteria for both perspectives.

Since this study aims at defining general applicable evaluation criteria for data and analytics use cases, the criterion of novelty was not taken into consideration as it is mainly relevant in the context of innovation initiatives.

Instead of market potential, some studies use different terms like profitability, benefit, financial outcome, or usefulness. All of these can be traced back to the value that results from a use case's outcome. For this reason, the first criterion that should be considered in the evaluation is the *added value* the use case is delivering.

If a use case is not aligned to the strategy and business of an organization, the results may not be considered as useful and valuable. This means that the use case may not be worthy of starting the development (Tauqeer & Bang, 2019, p. 6). Therefore, the *strategic fit* should also be included in the case of data and analytics.

Both criteria can be grouped together as they describe the *potential* of a use case from an organizational perspective.

As the feasibility cannot be determined that easily for data and analytics use cases, the study differentiates between *technical feasibility* and *economic feasibility* as a first step. Both are further divided into more specific and assessable criteria.

The *technical feasibility* specifically relates to the obstacles of data and analytics. As explained before, most data and analytics use cases do not make it into production. There are many studies that deal with the challenges and critical success factors as these are indicators that have an impact on the outcome of such a use case. Therefore, they were used to derive meaningful evaluation criteria to assess the feasibility of the use case.

In general, the challenges and critical success factors concerning data and analytics projects can be grouped into: *data, tools and technologies, people,* and *processes and management* (Al-Sai et al., 2019; Ermakova et al., 2021; Sivarajah et al., 2017). In this paper, the same categories are used as critical success factors are derived from the challenges (Al-Sai et al., 2020; Saltz & Shamshurin, 2016; Sim, 2014).

The data related challenges and critical success factors include issues concerning organizational data management processes (Al-Sai et al., 2020, p. 118953) and the data characteristics itself (Sim, 2014, p. 69).

Especially the data access and quality are considered as major impacts on the success of data-driven initiatives (Ermakova et al., 2021, p. 5086). Therefore, the evaluation criteria of *data availability and access*, and *data quality* are taken into consideration.

To develop data and analytics use cases, appropriate tools for collecting, storing, processing and analyzing the data are needed. Also, a solid and scalable IT architecture, and the right algorithms are fundamental (Al-Sai et al., 2020, p. 118953; Hilbert, 2005, p. 238; Sim, 2014, p. 70; Yeoh et al., 2008, p. 89). To cover this part, the evaluation criterion of *tools and technologies* is added.

Furthermore, a cross-functional and interdisciplinary project team composition with qualified people is essential (Sim, 2014, p. 70; Yeoh et al., 2008, p. 88). This includes hard skills like coding experience and familiarity with certain tools (Liu, 2019), and soft skills like communication, storytelling, creativity, and curiosity (Ismail & Zainal Abidin, 2016). Also, a certain business and domain knowledge is important to understand the business-related data and to be able to draw meaningful conclusions (Chen et al., 2012, p. 1183). All these skills are condensed in the evaluation criteria of *expertise*.

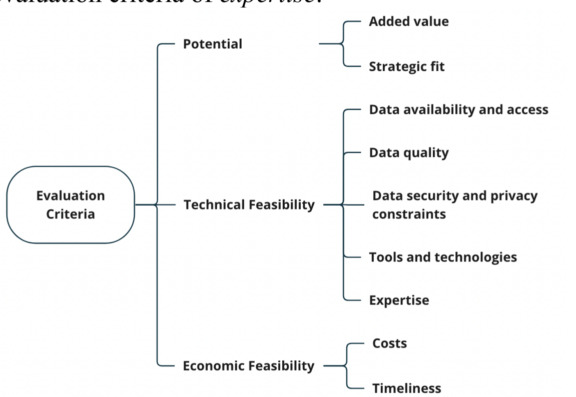

Figure 1. Evaluation criteria of data and analytics use cases.

The processes and management category includes a series of how things are done when dealing with data. Examples are a clear data strategy and defined processes. Also, general aspects like collaboration, communication, management support and an existing change management are relevant (Al-Sai et al., 2019, p. 154; Chen et al., 2006, p. 285; Cronholm et al., 2017, pp. 5–6; Hilbert, 2005, p. 238; Sivarajah et al., 2017; Yeoh et al., 2008, p. 86). Handling privacy and security issues in regard of data belongs to the same category. This is a major challenge for companies as those regulations can lead to a failure of data initiatives when not addressed correctly (Ermakova et al., 2021; Sivarajah et al., 2017, p. 12). As this aspect is also of highest importance for Bundesdruckerei GmbH, the criterion of *data security and privacy constraints* is assessed in the evaluation. The literature mentions time and financials as common evaluation criteria. Also, the lack of budget and time has a major impact on data and analytics use cases (Ermakova et al., 2021; Weiss, 2009). Therefore, *costs* and *timeliness* are part of the use case evaluation in order to determine the *economic feasibility*.

Figure 1 provides an overview of the criteria used in this paper for the evaluation of a data and analytics use case. The paper does not claim completeness in this matter as there might be additional relevant indicators.

4.2. Assessment of criteria

Firstly, the *added value* of the evaluated data use case is assessed. The value of a data and analytics use case can be distinguished by four potential dimensions: *value creation, value capture, value proposition,* or *value network* (Kayser et al., 2021). The value can be estimated qualitatively or quantitatively (Zolnowski et al., 2017, p. 187). The estimated value is the base to evaluate the criterion.

To evaluate the strategic fit, the use case should be analyzed in regard of its compatibility with aspects like corporate philosophy, image, policies, mission, objectives, character, or management interests (Karger, 1983, p. 46). Based on that, the use case's impact on the achievement of the corporate goals and strategy can be derived. It is also helpful to think about the consequences of rejecting a certain use case to determine its impact.

To address the criteria of *data availability and access*, the needed data and its sources have to be analyzed first. The knowledge about the data source supports the identification of people responsible for respective systems who might be important stakeholders (Kayser et al., 2019, p. 6). Humblet et al. (2016, p. 3) describe data availability as the "degree to which data can be instantly accessed". The access is determined by the sensitivity of the data. It can be distinguished between public/open, internal, confidential or restricted data (Kayser et al., 2019, p. 5). Based on that, the user can have free, limited, or no access to the data (Humblet et al., 2016, p. 4). A data gap analysis is a suitable instrument to assess this criterion. This includes the comparison between relevant data that is already available and accessible to the company, and data that is needed but lacking. It should also be considered if a different granularity, dimension, or frequency is required, and if enough data is available (UNDP & UN Global Pulse, 2016, p. 11).

The data quality is evaluated with the help of quality dimensions. Reasonable dimensions are completeness, timeliness, validity, understandability, consistency, or integrity (Sebastian-Coleman, 2013; Vetrò et al., 2016).

The sensitivity of the data does not only determine the data access, but also security and privacy constraints that influence the data processing and deployment of the solution (Baier et al., 2019). Especially legal frameworks like the *European General Data Protection Regulation (GDPR)* that relate to data privacy protection need to be considered. The trade-off is between the need of personal information for data analytics and the protection of individual identities (Malle et al., 2017, p. 155). Those legal constraints require appropriate security measures to deal with potential risks to affected individuals when processing identifiable personal information, biometric, or health data (Radley-Gardner et al., 2016). Therefore, the sensitivity level of the data is a suitable indicator to evaluate this criterion. Categories like the data security levels concerning research data used by Harvard University (Harvard Information Security, n.d.) are helpful to define the sensitivity level.

Similar to the data availability, a gap analysis is useful to assess tools and technologies that are required for developing the use case and to discover gaps. The same applies for the expertise. The more gaps there are in the internal skillset and the more tools and technologies are missing, the harder it is to develop the use case.

The cost estimation of data and analytics use cases is difficult, especially in the early project phase (Marbán et al., 2008, p. 134). The same is true for the estimation of the timeliness as those use cases include many unknowns (Geller, 2021). Marbán et al. (2007) summarize drivers that determine the effort of data and analytics use cases. Those drivers are related to data, models, platforms, tools, the project itself, and the staff. As effort determines the costs and timeliness of the use case, those drivers can be used for the evaluation of both criteria. Most drivers can be linked to the technical feasibility criteria. This means that the economic feasibility is determined by the technical feasibility.

5. Framework implementation

Atlassian Confluence was used as a tool to implement the framework within Bundesdruckerei GmbH. By that, the evaluation is directly documented and can be shared across the organization. Interactive elements offered by Atlassian Confluence are embedded. Surveys are used to enable a user-friendly rating and to foster collaboration by giving the possibility to include several people in the process. Additionally, a radar chart was added to visualize the scores for each criterion. Figure 2 shows the structure of the framework and a generalized overview of the content from each section.

The first section of the framework is about describing the use case. It includes an explanation of the problem, the solution, and the addressed user. At this part, the user of the framework is also asked to think about KPIs and related objectives that can be used to measure the success of the initiative. Furthermore, the initiator, the responsible, and the current status of the use is mentioned.

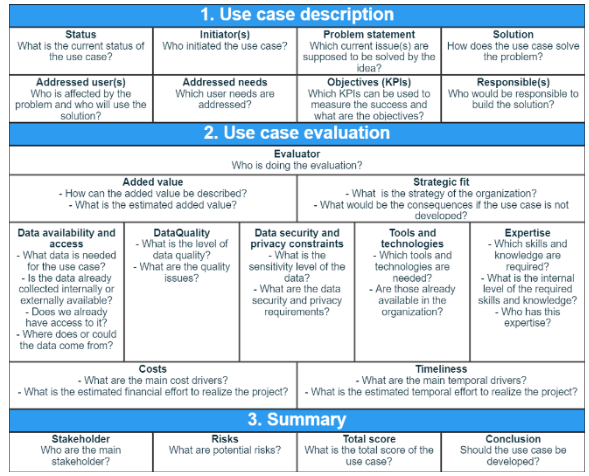

Figure 2. Framework for data and analytics use case evaluation

The main part of the framework is the evaluation section. It contains the evaluation criteria and corresponding evaluation questions to guide the user. Additional explanations and examples are provided to support the process. The evaluation is based on a scoring model. This means that each criterion receives a score which are summed up in the end. All use cases that exceed a minimum final score during the evaluation are selected (Baker & Albaum, 1986, p. 34). Scoring models are a popular way for a formal idea evaluation, since their usage, understanding, and communication to others are easy (Baker & Albaum, 1986, p. 34). It is also easily adaptable as weights can be included or criteria can be exchanged.

The scoring scale ranges from 1 to 5 points for each criterion. Each point on the scale is defined to increase the comparability of each evaluation. As there are nine criteria, the maximum score of 100% equals 45 points. Based on that result, the use case is rejected, revised, or accepted. If not accepted, the use case should be revised and further analyzed. Maybe it just needs to be adapted to get more feasible, or it could be put to an idea pool and re-evaluated at another time, when the required data, technologies, or expertise are available.

In the last section of the framework, the results are summarized and visualized by using a radar chart. This allows a quick assessment of the strengths and weaknesses of the use case. Furthermore, the relevant stakeholders and risks can be derived from the evaluation part. Based on the final score, a recommendation is made in the end. By that, several use cases can be compared and prioritized.

6. Evaluation

Two iterations were conducted to improve and evaluate the framework. The qualitative feedback and the quantitative data from the evaluation questionnaire were used as a base. In the following, the key findings of both iterations and the improvements that were made are summarized.

Key Findings. The summary of the key findings follows the structure of the criteria proposed: usefulness/effectiveness, usability, practicality, efficiency, and design.

In general, the framework was perceived as useful and effective regarding the evaluation and prioritization of data and analytics use cases. The elements and criteria of the framework were considered as relevant and important for the evaluation. In addition, one participant mentioned that it encourages the user to think thoroughly about the use case. Others commented that it provides a good overview of relevant aspects that need to be considered when making a deciding. Therefore, the framework is effective as the objective of enabling people to evaluate use cases and draw a conclusion was achieved. This was also confirmed by the quantitative data. Almost all participants strongly agreed that the framework supports the assessment of data and analytics use cases and that it enables a comparison and prioritization between them. Furthermore, the participants agreed that the framework facilitates communication between initiators of use cases and decision makers.

Regardless of the data literacy level of the participants, they outlined the clear structure of the framework. Especially, the introduction text, the interactive list of content and the summarizing radar chart were appreciated. Furthermore, the examples and scoring scales were considered as helpful. This also reflects in the quantitative results as all participants either agree or strongly agree that the framework is clearly structured, the instructions and examples are helpful, and the scales enable a clear scoring. In difference to that, there was a mixed feedback about the intuitiveness and self-explanation of the framework. While some participants had a neutral opinion about the corresponding statements, other strongly agreed.

The practicality was rated positively as the framework could be easily adapted and used by different departments and organizations according to the given feedback. But the participants had a diverse opinion about the implementability. While some participants had some doubts, others thought that it is easily implementable. This could mean that some participants considered the technical implementation while others based their decision on the organizational aspects like open-mindedness or management support.

All in all, the usage of the framework is evaluated as efficient. The participants see the invested time that is needed to fill out the framework (about three hours in the testing) as appropriate. Furthermore, the majority did not think that the same output quality could be reached with fewer criteria. One participant mentioned that "*the criteria are well chosen and important*". Only one person had a neutral point of view about that. While three people believed that a rapid consensus and decision making could be reached with the framework, two participants gave a "*neutral"* response.

The majority of the participants reflected the design as appealing. Also, most of them agreed that the design supports the usability of the framework. Just one participant had a neutral opinion about both statements.

The testing revealed that the analyzed use case suggested by the manufacturing department had some major weaknesses. The use cases received an overall score of 58% and therefore it was put into revision. On the one hand, the use cases added a low value and had a low strategic fit. On the other hand, the poor data quality and relatively high data security and privacy constraints negatively influenced the technical feasibility. This means that the effort for realizing the idea might not bring enough value to the company.

In general, no major problems were encountered while testing the framework. Questions were mainly directed on the meaning of specific words. Therefore, more explanations were added afterwards. Other smaller adjustments of scales, introduction texts, or guiding questions were made. The same applies for the second iteration.

The participants of both iterations had an overall positive opinion towards the use case evaluation framework. Especially the clear structure, the instructions, the examples, and the scales were appreciated. On the other hand, the participants gave more neutral feedback regarding the organizational implementability and the self-explanation of the framework elements. Also, some doubts existed if the framework enables a rapid decision making.

These aspects seem to hold the highest potential for improving the framework. Additional testing with different companies and participants is required.

7. Discussion

This paper presents criteria and a framework to evaluate data and analytics use cases at an early stage. It was developed in close cooperation with Bundesdruckerei GmbH. Following the action design research approach, a total of three iterative design cycles were carried out to build, adjust and evaluate the artifact. The iterations were accompanied by qualitative interviews and quantitative questionnaires to include diverse perspectives.

In general, the evaluation framework supports decision makers to identify weaknesses and strengths of data and analytics use cases at an early stage. The usage of the tool reduces the risk for an organization of investing into use cases that are unbeneficial or unfeasible. Therefore, the tool includes criteria that determine the potential, technical feasibility, and economic feasibility. They were derived from criteria used in general innovation management on the one hand and critical success factors concerning data and analytics use cases on the other hand.

In addition to that, the framework supports initiators to communicate their ideas for use cases to the corresponding decision makers. The underlying scoring model to evaluate the criteria facilitates the communication and the usability, as it is easy to understand and use. Also, the scoring model enables a comparison and prioritization of use cases. The evaluation criteria and scores could also be used to develop a model that predicts the probability of success of data and analytics use cases.

Limitations. This study reached its aim to develop a framework to evaluate data and analytics use cases at an early stage. However, the research has to be seen in the light of its limitations. The tool was developed in close cooperation with a small number of experts of one organization. To a certain extent, the design and content was influenced by Bundesdruckerei GmbH employees who participated in the iteration cycles. Further research would be needed to see how other companies would be able to implement the framework.

The framework was tested on a single use case. Also, the iterative design cycles were conducted with a limited number of participants. The small sample size limits the generalizability of the evaluation. A further testing on more ideas, in different environments, and with a larger number of participants is needed.

Another limitation comes with the challenge of an early stage evaluation. Although the possibility of making changes is the greatest at this stage, much must be based on assumptions (Samset & Christensen, 2017, p. 4). The lack of information often leads to reliance on experience, opinion, or even guesswork.

The framework is built on a scoring model that facilitates the usability, understandability, and transparency of the evaluation. But scoring has important shortcomings such as the use of criteria that are not independent of each other, the inconsistent application of criteria, difficulties with the interpretation of scores, or the assumption that raters have all the necessary knowledge (Baker & Albaum, 1986, p. 34).

Furthermore, the chosen evaluation criteria have not been empirically derived as they are based on literature research. This creates the risk that important criteria have been missed or other ones would be more meaningful in determining the quality of a data and analytics use case. A survey with seven experts was conducted to reduce this risk. The same problematic applies to the scales. They were developed with the aim of reducing the evaluation's subjectivity.

Also, the boundaries for accepting, revising, or rejecting an idea need to be tested on more examples to check if they are appropriate. All these limitations can influence the performance of the framework and reduce the credibility of the evaluation results.

Therefore, the framework should be additionally tested by a retrospective evaluation of several successful and unsuccessful data and analytics use cases. This would show, if the chosen criteria and the final result of the evaluation framework are really meaningful to evaluate a use case.

Future Work. Additional iterative design cycles with further testing and evaluations would be the next step in terms of the ADR approach. These should be done with different organizations and experts that do not work for Bundesdruckerei GmbH, to include new perspectives. This would show if the chosen criteria were perceived as suitable to determine the quality of a use case. Furthermore, it would be useful to analyze if the framework really facilitates a rapid decision making and reaching of consensus compared to the existing approaches of organizations.

It would also be interesting to use the framework to do a retrospective (*ex post*) evaluation of data and analytics use cases. If enough data is collected, the criteria could be empirically analyzed to show their importance in determining the probable success of a use case. Additional criteria that haven't been used in the framework could be included in this process. This could also be used as a base to develop different kinds of models to predict the success rate of data and analytics use cases.

8. References

Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2019). Big Data Impacts and Challenges: A Review. *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, 150–155.

Al-Sai, Z. A., Abdullah, R., & Husin, M. H. (2020). Critical Success Factors for Big Data: A Systematic Literature Review. *IEEE Access*, *8*, 118940–118956.

Baier, L., Jöhren, F., & Seebacher, S. (2019). *Challenges in the deployment and operation of machine learning in practice*.

1–15. *ECIS*
Baker, K. G., & Albaum, G. S. (1986). Modeling New Product Screening Decisions. *Journal of Product Innovation Management*, *3*(1), 32–39.
Carbonell-Foulquié, P., Munuera-Alemán, J. L., & Rodríguez-Escudero, A. I. (2004). Criteria employed for go/no-go decisions when developing successful highly innovative products. *Industrial Marketing Management*, *33*(4), 307–316.
Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). *CRISP-DM 1.0*. SPSS.
Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, *36*(4), 1165–1188.
Chen, Y., Hu, D., & Zhang, G. (2006). Data Mining and Critical Success Factors in Data Mining Projects. In K. Wang, G. L. Kovacs, M. Wozny, & M. Fang (Eds.), *Knowledge Enterprise: Intelligent Strategies in Product Design, Manufacturing, and Management* (Vol. 207, pp. 281–287). Springer US.
Ciriello, R. F., Richter, A., & Schwabe, G. (2016). Designing an Idea Screening Framework for Employee-Driven Innovation. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 4262–4271.
Colas, M., Finck, I., Buvat, J., Nambiar, R., Singh, R. R., & *Cracking the Data Conundrum: How Successful Companies Make Big Data Operational* (2014). Digital Transformation Research Institute.
Cooper, R. G. (1981). An empirically derived new product project selection model. *IEEE Transactions on Engineering Management*, *EM-28*(3), 54–61.
Cooper, R. G. (1990). Stage-gate systems: A new tool for managing new products. *Business Horizons*, *33*(3), 44–54. Business Source Ultimate.
Cooper, R. G. (2011). Perspective: The Innovation Dilemma: How to Innovate When the Market Is Mature: HOW TO INNOVATE WHEN THE MARKET IS MATURE. *Journal of Product Innovation Management*, *28*(s1), 2–27.
Cooper, R. G., & De Brentani, U. (1984). Criteria for screening new industrial products. *Industrial Marketing Management*, *13*(3), 149–156.
Cronholm, S., Göbel, H., & Rittgen, P. (2017). Challenges Concerning Data-Driven Innovation. *ACIS 2017 Proceedings*, *3*, 1–12.
Dean, D., Hender, J., Rodgers, T., & Santanen, E. (2006). Identifying Quality, Novel, and Creative Ideas: Constructs and Scales for Idea Evaluation. *Journal of the Association for Information Systems*, *7*(10), 646–699.
Ermakova, T., Blume, J., Fabian, B., Fomenko, E., Berlin, M., & Hauswirth, M. (2021). Beyond the Hype: Why Do Data-Driven Projects Fail? *Proceedings of 54th Hawaii International Conference on System Sciences*, 5081–5090.
Geller, A. (2021, June 11). 10 Reasons Why Estimating Time For Data Projects is Hard. *GoodData Developers*. https://medium.com/gooddata-developers/10-reasons-why-estimating-time-for-data-projects-is-hard-218a202ba7e4
Gerlach, S., & Brem, A. (2017). Idea management revisited: A review of the literature and guide for implementation. *International Journal of Innovation Studies*, *1*(2), 144–161.
Hammedi, W., van Riel, A. C. R., & Sasovova, Z. (2011). Antecedents and Consequences of Reflexivity in New Product Idea Screening*: Antecedents and Consequences of Reflexivity in New Product Idea Screening. *Journal of Product Innovation Management*, 662–679. h
Hansen, M. T., & Birkinshaw, J. (2007). The Innovation Value Chain. *Harvard Business Review*.
Harvard Information Security. (n.d.). *Data Security Levels—Research Data Examples*. Harvard University. Retrieved November 12, 2021, from https://security.harvard.edu/data-security-levels-research-data-examples
Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75–105.
Hilbert, A. (2005). Critical Success Factors for Data Mining Projects. In D. Baier, R. Decker, & L. Schmidt-Thieme (Eds.), *Data Analysis and Decision Support* (pp. 231–240). Springer-Verlag.
Huang, Z., Ahmed, C., & Mickael, G. (2020). A model for supporting the ideas screening during front end of the innovation process based on combination of methods of EcaTRIZ, AHP, and SWOT. *Concurrent Engineering*, *28*(2), 89–96.
Humblet, M.-F., Vandeputte, S., Mignot, C., Bellet, C., De Koeijer, A., Swanenburg, M., Afonso, A., Sanaa, M., & Saegerman, C. (2016). How to Assess Data Availability, Accessibility and Format for Risk Analysis? *Transboundary and Emerging Diseases*, *63*(6), e173–e186.
Ismail, N. A., & Zainal Abidin, W. (2016). Data Scientist Skills. *IOSR Journal of Mobile Computing & Application*, *03*(04), 52–61.
Karger, T. (1983). Screening new product/marketing ideas to reduce the chance of making expensive mistakes. *Management Review*, *72*(8), 45–48. Business Source Ultimate.
Kayser, L., Fruhwirth, M., & Mueller, R. M. (2021). Realizing Value with Data and Analytics: A Structured Literature Review on Classification Approaches of Data-Driven Innovations. *Proceedings of 54th Hawaii International Conference on System Sciences*, 5686–5695.
Kayser, L., Mueller, R. M., & Kronsbein, T. (2019). *Data Collection Map: A Canvas for Shared Data Awareness in Data-Driven Innovation Projects*. 18.
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.
Kim, J., & Wilemon, D. (2002). Strategic issues in managing innovation's fuzzy front-end. *European Journal of Innovation Management*, *5*(1), 27–39.
Kim, W. C., & Mauborgne, R. (2000). Knowing a Winning Business Idea When You See One. *HARVARD BUSINESS REVIEW*, *78*(5), 129–138.
Kronsbein, T., & Mueller, R. M. (2019). Data Thinking: A Canvas for Data-Driven Ideation Workshops. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 561–570.
Kudrowitz, B. M., & Wallace, D. (2013). Assessing the quality of ideas from prolific, early-stage product ideation. *Journal of Engineering Design*, *24*(2), 120–139.
Kühne, B., & Böhmann, T. (2020). Formative Evaluation of Data-Driven Business Models – The Data Insight Generator. *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 427–436.
Lauto, G., & Valentin, F. (2016). How preference markets assist new product idea screening. *Industrial Management & Data Systems*, *116*(3), 603–619.
Liu, S. (2019). *Most wanted data science skills U.S. 2019*. Statista. https://www.statista.com/statistics/1016247/united-states-wanted-data-science-skills/
Magnusson, P. R., Netz, J., & Wästlund, E. (2014). Exploring

holistic intuitive idea screening in the light of formal criteria. *Technovation*, *34*(5–6), 315–326.

Malle, B., Kieseberg, P., & Holzinger, A. (2017). DO NOT DISTURB? Classifier Behavior on Perturbed Datasets. In A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), *Machine Learning and Knowledge Extraction* (Vol. 10410, pp. 155–173). Springer International Publishing.

Marbán, O., Menasalvas, E., & Fernández-Baizán, C. (2008). A cost model to estimate the effort of data mining projects (DMCoMo). *Information Systems*, *33*(1), 133–150. h

Moon, H., & Han, S. H. (2016). A creative idea generation methodology by future envisioning from the user experience perspective. *International Journal of Industrial Ergonomics*, *56*, 84–96.

Neifer, T., Lawo, D., & Esau, M. (2021). Data Science Canvas: Evaluation of a Tool to Manage Data Science Projects. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 5399–5408. h

NewVantage Partners. (2021, January 4). *NewVantage Partners Releases 2021 Big Data and AI Executive Survey, The Journey to Becoming Data-Driven, A Progress Report in the State of Corporate Data Initiatives*. https://www.businesswire.com/news/home/20210104005022/en/NewVantage-Partners-Releases-2021-Big-Data-and-AI-Executive-Survey

Radley-Gardner, O., Beale, H., & Zimmermann, R. (Eds.). (2016). *Fundamental Texts On European Private Law*. Hart Publishing.

Rochford, L. (1991). Generating and Screening New Product Ideas. *Industrial Marketing Management*, *20*, 287–296.

Salerno, M. S., Gomes, L. A. de V., Silva, D. O. da, Bagno, R. B., & Freitas, S. L. T. U. (2015). Innovation processes: Which process for which project? *Technovation*, *35*, 59–70.

Saltz, J. S., & Shamshurin, I. (2016). Big data team process methodologies: A literature review and the identification of key factors for a project's success. *2016 IEEE International Conference on Big Data (Big Data)*, 2872–2879. h

Samset, K., & Christensen, T. (2017). Ex Ante Project Evaluation and the Complexity of Early Decision-Making. *Public Organization Review*, *17*(1), 1–17.

Sandström, C., & Björk, J. (2010). Idea management systems for a changing innovation landscape. *International Journal of Product Development*, *11*(3/4).

Schmidt, J. B., Sarangee, K. R., & Montoya, M. M. (2009). Exploring New Product Development Project Review Practices. *Journal of Product Innovation Management*, *26*, 520–535.

Sebastian-Coleman, L. (2013). Data Quality and Measurement. In *Measuring Data Quality for Ongoing Improvement* (pp. 39–53). Elsevier.

Sein, Henfridsson, Purao, Rossi, & Lindgren. (2011). Action Design Research. *MIS Quarterly*, *35*(1), 37. h

Sim, J. (2014). Consolidation of Success Factors in Data Mining Projects. *GSTF Journal on Computing*, *4*(1), 66–73.

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.

Stevanovic, M., Marjanovic, D., & Storga, M. (2015). A MODEL OF IDEA EVALUATION AND SELECTION FOR PRODUCT INNOVATION. *Proceedings of the 20th International Conference on Engineering Design (ICED15)*, *nn*, 10.

Tab, M., & Sharkey, K. (2020). *What is the Team Data Science Process? - Azure Architecture Center*. Microsoft. https://docs.microsoft.com/en-us/azure/architecture/data-science-process/overview

Tauqeer, M. A., & Bang, E. (2019). A tool for idea screening by assortment of existing literature. *Proceedings of ISPIM Conferences*, 1–9. Business Source Ultimate.

Toubia, O., & Florès, L. (2007). Adaptive Idea Screening Using Consumers. *Marketing Science*, *26*(3), 342–360.

UNDP, & UN Global Pulse. (2016). *A Guide to Data Innovation for Development: From Idea to Proof of Concept*. https://www.undp.org/publications/guide-data-innovation-development-idea-proof-concept

Vanauer, M., Bohle, C., & Hellingrath, B. (2015). Guiding the Introduction of Big Data in Organizations: A Methodology with Business- and Data-Driven Ideation and Enterprise Architecture Management-Based Implementation. *2015 48th Hawaii International Conference on System Sciences*, 908–917.

VentureBeat. (2019, July 19). Why do 87% of data science projects never make it into production? *The Machine*. https://venturebeat.com/2019/07/19/why-do-87-of-data-science-projects-never-make-it-into-production/

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, *33*(2), 325–337.

Weiss, G. M. (2009). *Data Mining in the Real World: Experiences, Challenges, and Recommendations*. 8.

White, A. (2019, January 3). *Our Top Data and Analytics Predicts for 2019*. https://blogs.gartner.com/andrew_white/2019/01/03/our-top-data-and-analytics-predicts-for-2019/

Yarmohammadi, M., Rezvani, M., & Alborzi, M. (2017). How Complementors Screen new Product Ideas: A Qualitative Multiple Case Study. *European Conference on Innovation and Entrepreneurship*. 669–676.

Yeoh, W., Koronios, A., & Gao, J. (2008). A Critical Success Factors Framework. *International Journal of Enterprise Information Systems*, *4*(3), 79–94.

Zawadzki, J. (2020, January 6). *Introducing the AI Project Canvas*. Towards Data Science. https://towardsdatascience.com/introducing-the-ai-project-canvas-e88e29eb7024

Zolnowski, A., Anke, J., & Gudat, J. (2017). Towards a Cost-Benefit-Analysis of Data-Driven Business Models. *Proceedings Der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*, 181–195.