6-15-2020

# A STATE-OF-THE-ART OVERVIEW AND FUTURE RESEARCH AVENUES OF SELF-SERVICE BUSINESS INTELLIGENCE AND ANALYTICS

Sven Michalczyk
*Karlsruhe Institute of Technology (KIT)*, sven.michalczyk@kit.edu

Mario Nadj
*Karlsruhe Institute of Technology (KIT)*, mario.nadj@kit.edu

Darius Azarfar
*NONE*, darius.azarfar@gmail.com

Alexander Maedche
*Karlsruhe Institute of Technology (KIT)*, alexander.maedche@kit.edu

Christoph Gröger
*Robert Bosch GmbH*, christoph.groeger@de.bosch.com

# A STATE-OF-THE-ART OVERVIEW AND FUTURE RESEARCH AVENUES OF SELF-SERVICE BUSINESS INTELLIGENCE AND ANALYTICS

*Completed Research Paper*

Michalczyk, Sven, Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, sven.michalczyk@kit.edu

Dr. Nadj, Mario, Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, mario.nadj@kit.edu

Azarfar, Darius, ITK Engineering GmbH, Rülzheim, Germany, darius.azarfar@itk-engineering.de

Professor Dr. Mädche, Alexander, Institute of Information Systems and Marketing (IISM), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany,

Dr. Gröger, Christoph, Robert Bosch GmbH, Stuttgart, Germany, christoph.groeger@de.bosch.com

## Abstract

*Self-Service Business Intelligence and Analytics (SSBIA) is an upcoming approach and trend that enables casual business users to prepare and analyze data with easy-to-use Business Intelligence and Analytics (BIA) systems without being reliant on expert support or power users to perform their (complex) analytical tasks easier and faster than before. Despite a strong interest of scholars and practitioners in SSBIA, the understanding about its underlying characteristics is limited. Furthermore, there is a lack of a structured and systematic form in which SSBIA research can be classified. Against this backdrop, this article showcases the current state-of-the-art of SSBIA research along four key areas in the field: (1) perspectives on SSBIA, (2) user roles involved, (3) required expertise, and (4) supported levels of self-service. Analyzing 60 articles, our main contribution resides in the synopsis of SSBIA literature in these four areas. For instance, we illustrate that there exist three perspectives of SSBIA: artefact-centric (45% of analyzed studies), user-centric (82%), and governance-centric (25%). On the basis of our analysis, we suggest promising avenues, which will support scholars in their endeavors on how to pursue with future avenues in the field of SSBIA (for e.g., understanding the trade-off between top-down and bottom-up capabilities).*

*Keywords: Self-Service, Business Intelligence, Analytics, Literature Review*

## 1 Introduction

Business Intelligence and Analytics (BIA) refers to all "techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions" (Chen et al., 2012, p. 1166). Logi (2017) illustrates that 66% of organizations use some sort of BIA tool to drive their business, whereas only 6% of organizations do not consider applying such systems. The size of the global BIA market is also expected to grow from $15.64 billion in 2016 to reach $29.48 billion by 2022 with a compound annual growth rate of 11.1% (BARC, 2019). Furthermore, as the provisioning of the right information to the right person at the right time is critical to remain competitive and represents a key concern of BIA (Bucher et al., 2009), new solutions such as data lakes – “a set of one or more data repositories that have been created to support data discovery, analytics, ad hoc investigations, and reporting” (Chessell et al., 2014; p. 1) – are still high in demand (e.g., Lennerholt et al., 2018).

A prominent example case refers to enable casual business users (referred in the following as casual users) to work with data by collaborating with skilled power or IT users (Alpar and Schulz, 2016). For instance, research suggests developing integrated collaboration platforms within the BIA landscape to support knowledge sharing and to solve complex issues with diverse user groups (e.g., Bani-Hani et al., 2018a; Passlick et al., 2017). Hereby, organizations face the challenge to support different user roles with diverse levels of expertise for various analytical demands ranging from information usage (e.g., access to reports) over information creation (e.g., creation of reports) towards information resources creation (e.g., harnessing of new data sources) (Alpar and Schulz, 2016). The common denominator to address these challenges refers to a branch of BIA technologies summarized under the term Self-Service Business Intelligence and Analytics (SSBIA). A major idea of SSBIA is that casual users will be able to prepare and analyze data with easy-to-use BIA systems without the need for expert support (Lennerholt et al., 2018). Well-known tools, such as KNIME (2019) and Tableau (2019), already offer self-service capabilities. KNIME represents a tool for interactive data analysis with the ability to implement numerous modules of Machine Learning (ML) methods (KNIME, 2019). Hereby, ML refers to the application of algorithms that improve with experience and learn from data (Mitchell, 1997). In turn, Tableau is an interactive visualization tool providing users the opportunity to link graphical user interfaces with drag and drop (Tableau, 2019). In general, SSBIA can empower causal users to create and share reports on their own from high volumes of diverse data without getting stuck into long running business requests with IT departments. Or as Eckerson (2009) has summed it up: “Users get exactly the reports they want, when they want (…)” (p. 2). In addition, SSBIA can help power users to perform their (complex) analytical tasks easier and faster than before (Alpar and Schulz, 2016).

In defiance of the relevance of SSBIA, a systematic approach to illustrate the current state-of-the-art in this domain along the various SSBIA characteristics is missing to offer suggestions for further research avenues. In addition, despite the relatively young nature of the SSBIA field (e.g., Bani-Hani et al., 2018a), numerous articles have already been published, particularly in the last decade (e.g., Saeed and Abdinnour, 2011; Spahn et al., 2008). As a result, analyzing this area in a systematic way is required owing to several reasons: First, despite the increase of SSBIA systems, studies have shown that their adoption is still challenging due to the dependency on technical or organizational boundaries demanding common consensus about how to design and embed these systems in organizations. Following Logi (2017), only 21% of the respondents have data access when required illustrating that implementing SSBIA systems is not an easy endeavor. In this light, Lennerholt et al. (2018) emphasize the importance of addressing the complexity within SSBIA infrastructures as well as considering organizational boundaries for data management. Our analysis supports this debate by offering an overview of SSBIA characteristics which need to be considered when adopting such systems. Second, the vast number of studies requires structuration. A structure offers assistance in classifying and delimiting the research results by relevant properties. On this basis, we present a state-of-the-art overview and identify future research avenues (Kitchenham and Charters, 2007). Third, although some scholars sketched out SSBIA challenges (e.g., Lennerholt et al., 2018; Lismont et al., 2019; Sulaiman and Gómez, 2018), future research needs more guidance to be well-directed.

Against this backdrop, our review structures systematically the scattered research results, analyzing 60 articles that explicitly refer to SSBIA. Our goal is to review the state-of-the-art of SSBIA research to recommend avenues for future research. Therefore, the following research question is formulated: *What are the potential avenues in Self-Service Business Intelligence and Analytics for future research?* In the following, we describe the foundations of our article in section two. In section three, we discuss the method of our systematic literature review (SLR). The results of our study are presented in section four. In section five, we outline future research directions. Section six concludes our article.

## 2 Foundations

Transforming raw data into actions includes three types of BIA technologies: (1) data warehouses (DWH) or data lakes, (2) analytical tools, and (3) reporting tools (Miloslavskaya et al., 2016; Watson, 2009). Operational systems and external sources offer data of varying quality, format, and meanings (Abelló et al., 2013). Next, source data can be loaded via ETL (Extract, Transform, Load) or ELT (Extract, Load, Transform) processes into a DWH or data lake. When the data are stored, tools that apply advanced analytics (Bose, 2009), such as KNIME or Tableau, are used to offer new insights before reports and dashboards are generated as a basis for decision-making and next actions (Trigo et al., 2014).

Hereby, different types of user roles (i.e., casual, power, and IT user) exist to realize a BIA system (Lennerholt et al., 2018). Power users have expert knowledge and are capable of conducting complex analytical tasks (De Mauro et al., 2018). For instance, a power user has the ability to find appropriate data from diverse source systems and varying formats to create a report. In turn, casual users such as managers or business users have less experience and commonly rely on support from power users to generate reports as a foundation for their decision-making process. However, change requests or misunderstandings for reports (e.g., requiring the inclusion of another key performance indicator) might result in a time-consuming exchange process between casual and power users. As nowadays data are created at large scale and stored in a DWH or a data lake, and BIA systems are used in general by an increasing number of employees (Llave, 2018), IT departments face the challenge to offer an efficient technical infrastructure to support these exchange processes between casual and power users (Alpar and Schulz, 2016). Bottlenecks within these processes (e.g., technical problems, usability, requirements' misunderstandings, or inability to generate reports on time) might impede timely decisions or even lead to situations where casual users make decision without considering (all) the present underlying data within a report (Imhoff and White, 2011; Lennerholt et al., 2018). Due to the increasing amount of data and share of employees who rely on some sort of BIA tool, SSBIA represents one upcoming approach and trend in the field of BIA (e.g., Bani-Hani et al., 2017). Following Lennerholt et al. (2018), users "should be able to access and query data, use predefined reports, analyze data or create their own reports, in order to make decisions on time" (p. 5056). Hereby, different levels of self-service can be distinguished ranging from information usage (e.g., access to reports – lowest self-service level) over information creation (e.g., creation of reports – intermediate self-service level) to information resources creation (e.g., harnessing of new data sources – highest self-service level) (Alpar and Schulz, 2016).

In summary, the purpose of SSBIA is to support the self-reliance of users, whereby an increased level of self-service corresponds to a higher level of self-reliance. As our article showcases the current state-of-the-art along four key areas including studies that serve different self-service levels, this piece of research could thus represent a cornerstone of further efforts to refine and expand the concept of SSBIA.

## 3 Method

Following the guidelines by Webster and Watson (2002) and Kitchenham and Charters (2007), we conducted a SLR that consists of three stages: (1) plan, (2) conduct, and (3) report (see Figure 1). In the plan stage, we identified the SLR need and created a review protocol to evaluate it. In the conduct stage, we executed database searches to analyze relevant studies. In the report stage, we illustrated our results.

Our search string was developed in several steps. An initial exploratory search was conducted on Google Scholar using the search term "*Self-Service* AND (*Business Intelligence* OR *Analytics*)". By reviewing the results and after several iterations, the final search string consisted of five parts: First, we extracted the foundational terms of (a) "*Self-Service*", (b) "*Business Intelligence*", and (c) "*Analytics*" from our research question. Regarding the Business Intelligence component, we also used a related term, namely "*Decision Support Systems*", and the general term "*Information Systems*".

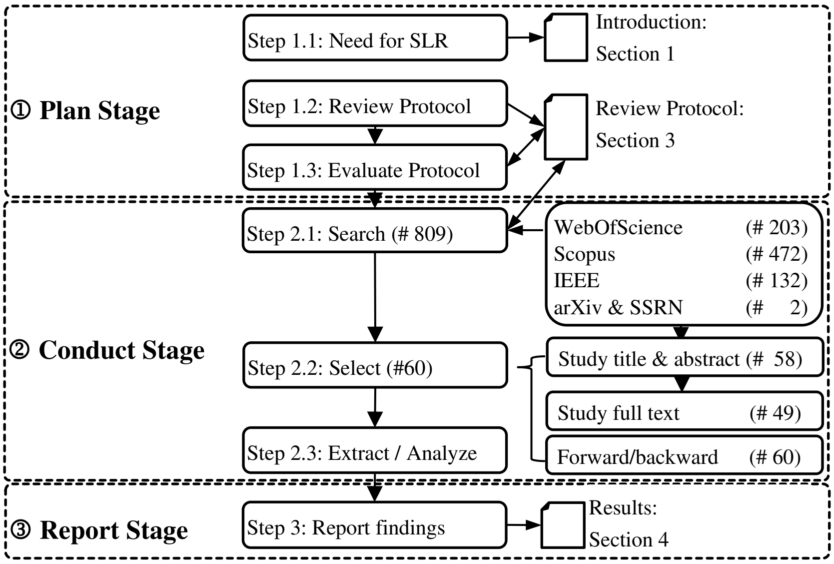


*Figure 1. Stages of the SLR*

Next, we included the term (d) "*Machine Learning*" as SSBIA solutions typically offer the possibility to integrate numerous modules of ML methods (e.g., KNIME). Following Alpar and Schulz (2016), the creation or usage of information respectively information resources involves the understanding of data. Therefore, we add various synonyms of data understanding such as data exploration and discovery. Thus, (e) all three terms (i.e., *data understanding*, *exploration*, and *discovery*) were used for our search string. Furthermore, we included the term data *visualization* because upcoming SSBIA tools such as Tableau use various forms of interactive visualizations for data exploration. The terms data *preparation*, data *wrangling*, and data *cleaning* were also used for our search string as they represent important activities for sharpening the data understanding and preparing the data for further analysis (Chapman et al., 2000; Llave, 2018). Within data preparation, we further identified the term data *integration* to be of relevance for our search string as it represents an important task to combine datasets from different sources, especially from relational databases. In addition, we also included the terms data *modeling* and data *mining* as self-service data mining tends to be an upcoming research interest (e.g., Zehnder and Riemer, 2018; Zschech et al., 2019). Finally, we used Boolean-operators and wildcards to create the final search string:

*"Self*Service" AND (("Business Intelligence" OR "Information System*" OR "Decision Support System*") OR "Analytic*" OR "Machine Learning" OR ("Data" AND ("Understanding" OR "Cleaning" OR "Preparation" OR "Integration" OR "Wrangling" OR "Modeling" OR "Mining" OR "Discovery" OR "Exploration" OR "Visualization")))*

Next, we selected WebOfScience (WoS), IEEE Library and Scopus for our SLR as these databases are well established and used by researchers as reliable sources for literature reviews (e.g., Bandara et al., 2015; Bondarouk et al., 2012; Hamari et al., 2014). To offer a holistic overview, we have not limited our review by year or publication outlet. On these grounds, we used the defined search string and found 203 studies for WoS, 132 for IEEE, and 472 for Scopus (807 in total). After the databases search, we conducted a hand search to reduce a potential source bias and also to consider unpublished studies by relying on ArXiv and SSRN (two studies were added). In the selection process, we manually excluded 751 studies by carefully scanning the title-, abstract-, and keyword-section and by applying the following selection criteria: Studies without BIA focus were dismissed. In particular, a BIA focus was given if the study at hand dealt with casual, power, or IT users supporting them to transform raw data into actions. Furthermore, we excluded studies that concentrated on general self-service technologies targeting customers, such as ATM machines. On this basis, 58 articles were left. Following the same criteria for a full text review, 49 relevant articles remained. Finally, we employed a forward and backward search and included another eleven studies to the sample (in total 60).

We applied a three-step approach to classify our results. In the **first step**, the codes for the categories were derived deductively from existing SSBIA literature to set up a solid basis for the classification. We followed the recommendations by Alpar and Schulz (2016) and De Mauro et al. (2018) and relied on codes for casual users (e.g., manager, business user), power users (i.e., data analyst, data scientist) and IT users (e.g., engineer, developer). Along this line, we also coded the required user experience as the purpose of a SSBIA system might strongly differ depending on whether it is used by novices or BIA experts (Lismont et al., 2019). Lastly, we relied on the concept proposed by Alpar and Schulz (2016) to investigate different levels of self-service and associated analytical tasks, namely, information usage as lowest self-service level (i.e., access to reports, access to drill anywhere reporting), information creation as intermediate self-service level (i.e., creation of reports, access to analytical functions), and information resources creation as highest self-service level (i.e., creation of mashups, harnessing of new data sources). Both the self-service levels and corresponding tasks were coded. The study of the different levels of self-service seems necessary as these levels shed light on how self-reliant users can perform their tasks and what degree of system support is required. In the **second step**, we used an inductive approach either to cross-validate the deductive codes or to create new codes. In particular, during the classification process, we added the task "deployment of analytical functions" to the self-service level "information resources creation" to allow for a more fine-granular investigation process. In addition, it became apparent that SSBIA encompasses several perspectives in the form of artifact-centric, user-centric, and governance-centric. In the **third step,** the studies were assigned independently to the respective codes by two researchers. Any discrepancy between the two coders' results was discussed and resolved by consulting a third researcher. In particular, mismatches occurred in the user role category as the assignment regarding the target user was sometimes not fully clear. We resolved the issue of mismatches by specifying that the target user of the SSBIA system does not only benefit from it but also uses the system.

# 4 Results

We classified the content of each study along the four categories: (1) perspectives, (2) user roles, (3) experience, and (4) levels of self-service. The classification is not mutually exclusive because the respective study might refer to several (sub-) categories. Thus, we created a concept matrix in the style as suggested by Webster and Watson (2002) (see Table 1). This section investigates these four categories as well as emphasizes the most prominent studies within each category.

## 4.1 Perspectives

In general, SSBIA is discussed along three perspectives in literature: (1) technical solutions (*artefact-centric view,* 45% of the studies), (2) drivers for user acceptance (*user-centric view,* 82%), as well as (3) governance challenges (*governance-centric view,* 25%).

### 4.1.1 Artefact-centric View

The artefact-centric perspective emphasizes that "the search, extraction, and integration of situational data should be accomplished by users through a continuous interaction with the application, without any mediation or intervention by analysts, designers, or programmers" (Abelló et al., 2013, p. 1). The underlying rational is to support users to make well-informed decisions by helping them to navigate through data via technical solutions. In Abelló et al. (2013), so called fusion cubes are proposed as an extension to data cubes by integrating situational data, which mostly refers to unstructured data from unspecified data sources that are not included in the DWH. A search-engine complements the approach enabling external data source discovery and selection, as well as on-demand ETL. The idea of the approach is to integrate fusion cubes into existing DWH frameworks already used within organizations. In turn, Chouder et al. (2017) propose an interactive approach for finding multidimensional structures in documents to enable querying. The structures are based on automatically mining through learning from past queries and approximate functional dependencies from data. In addition, Spahn et al. (2008) present an ontology-based architecture for a semantic query designer tool for casual users. The tool enables navigation and query building for business objects in reporting tasks. To automatically extract multidimensional elements from query logs, a framework is proposed by Ibáñez et al. (2017). With this information, a knowledge base is created to detect semantic correctness of categories by summarizing queries and their elements. Zehnder and Riemer (2018) involve multiple roles and tasks with the objective to empower casual users to leverage ML modules that have been specified by a domain-specific language.

### 4.1.2 User-Centric View

In turn, Imhoff and White (2011) do not only refer to technical aspects of SSBIA, but also emphasize a user-centric perspective with the aim to increase the user acceptance of SSBIA systems. Hereby, the authors propose four main objectives for achieving this aim: “easier access to source data for reporting and analysis, easier and improved support for data analysis features, faster deployment options such as appliances and cloud computing, and simpler, customizable, and collaborative end-user interfaces” (Imhoff and White, 2011, p. 5). Especially negative effects occur for casual users with scarce knowledge on how to apply SSBIA systems. For instance, Johannessen and Fuglseth (2016) showed that data modelling knowledge directly contributes to the consistency of the developed data cube and to the validity of the data analysis. However, the effect of learning on novice users is still an avenue for future research (Lismont et al., 2019). Against this backdrop, Convertino and Echenique (2017) supports the idea of implementing a multi-tool, multi-user role platform to create synergies between casual and power users. The underlying rational of the multi-tool and multi-user approach refers to integrating tools on a central platform for supporting various user groups within an organization for their daily tasks. In addition, visualization represents an important success factor (e.g., Imhoff and White, 2011 and Lousa et al., 2019 for the usage of information; Li et al., 2017 and Stodder, 2015 for the creation of information; Behringer et al., 2018 and Clark et al., 2016 for the creation of information resources). Hereby, knowledge from reference disciplines such as information visualization (e.g., Alpar and Schulz, 2016) or visual analytics (e.g., Behringer et al., 2018) are often times used as a backbone. For instance, Ahmed et al. (2013) suggest a taxonomy for BIA mash-ups, that “facilitate problem-solving in ad-hoc situational BI scenarios” without raising requirements to IT. In general, most studies emphasize the need to create easy-to-use graphical user interfaces (e.g., Imhoff and White, 2011) or to rely on interactive visualization techniques (e.g., Behringer et al., 2018) to increase user acceptance of SSBIA systems. In this line, Daradkeh and Al-Dwairi (2017) conclude that the quality of information, analysis and systems are key antecedents for perceived usefulness and ease of use of the SSBIA system at hand, and therefore strong predictors of user acceptance.

### 4.1.3 Governance-Centric View

Traditionally, a DWH represents a central database in organizations which is well-governed by IT departments. However, new data repositories with structured and unstructured data, typically called data lakes, are an important and upcoming direct source for SSBIA (Llave, 2018). SSBIA users should therefore also be allowed to access this data source which is typically not only IT-curated (Vo et al., 2018). However, errors could occur in the decision-making process because the data quality for these new data sources may not be suitable for the required analytical task (Stodder, 2015). Therefore, the adoption of data management principles, such as metadata concepts and data governance play a more important role within self-service data preparation (e.g., Imhoff and White, 2011; Passlick et al., 2017). However, there is a risk that, if these principles are too restrictive, business departments will start to build their own shadow IT (Stodder, 2015). In this regard, Imhoff and White (2011) prosed a rating and annotation system to determine the value of reports or their components. Such course of actions would allow the curation of reports to be sourced to the crowd, while IT users with insufficient knowledge of the business domain would no longer be responsible for assessing the quality of the reports or their components. In summary, we identified a trade-off in the area of governing SSBIA between a top-down environment (e.g., Mayer et al., 2015; Schuff et al., 2018;) that includes a pre-defined technical infrastructure, which typically lacks flexibility, and a bottom up environment (e.g., Abelló et al., 2013; Savinov, 2014) that gives users the full flexibly with the risk of taking wrong assumptions.

## 4.2 User Roles

SSBIA can address diverse types of user roles. Alpar and Schulz (2016) and De Mauro et al. (2018) distinguish between three user roles: (1) *casual* (77% of the studies, e.g., *managers*, *business users*), (2) *power* (i.e., *data analysts,* 68%; *data scientists,* 37%) and (3) *IT users* (18%, e.g., *engineers*, *developers*).

### 4.2.1 Casual Users

Casual users are associated with scarce knowledge for conducting analytical tasks. They typically obtain access to information already created or only need to specify parameters before processing them. In addition, they have access to reports and dashboard with a ‘‘drill anywhere’’ possibility (Alpar and Schulz, 2016). Self-reliant casual users are an important factor for the success of BIA tools (Lennerholt et al., 2018) as they have extensive knowledge of their business or engineering domain (Spahn et al., 2008). In this light, Englmeier and Román (2014) propose an ontology that enables casual users to learn from BIA-related knowledge. They claim that, over time, casual users naturally gather this knowledge and thereby also acquire a certain level of IT knowledge. Bilalli et al. (2016) conclude that user assistance is reachable by providing classifications for metadata in knowledge discovery processes. In turn, Varga et al. (2018) present a metadata framework to support user assistance to enable process automation by tracking all the processes casual and power users perform in transforming data into actions. Sessions, preferences, queries, and profiles are analyzed to automate the underlying process steps. Marjanovic (2015) takes a different view by defining consumers as a target group which seems promising as their analytical demands share similarities with casual users. With the proposed solution, every Australian citizen can analyze the performance of schools and students over time. The analysis is facilitated by visuals such as graphical comparisons and geographical maps, as well as functionalities for individual or collective insight-sharing.

### 4.2.2 Power Users

Power users cannot only access information but also create new information or information sources. Typical tasks focus on data integrity (e.g., cleaning data, or aggregating data sources) or analyzing prepared data (e.g., identifying outliers, applying ML algorithms) (Dinsmore, 2016). Common examples of power users refer to data analysts and data scientists (De Mauro et al., 2018).

In general, data scientists have a higher level of analytical and technical skills than data analysts and a deeper understanding of how data can be transformed into actions (De Mauro et al., 2018). This process often includes exploring structured and unstructured data to exploit new insights (Abelló et al., 2013). In particular, data scientists often create new insights by applying advanced analytics. Hereby, they also rely on other data sources such as external data or data from other organizational departments to extend the data base at hand (Abelló et al., 2013). Other than data analysts, data scientists are not necessarily bound to one department but can function across the organization at the cost of not having much insight of the underlying operational tasks (Eckerson, 2012). In contrast, data analysts are driven by the business impact they can make and are typically responsible for providing insights for casual users (Eckerson, 2012). For instance, they prepare data for diverse business teams in different departments of an organization. Their profound project management competence indicates that they work closely together with other user roles to achieve the defined business goals. Consequently, they share a higher business domain knowledge compared to data scientists (De Mauro et al., 2017). SSBIA systems support different user roles to carry out their analytical tasks on their own. Such system support compensates for the missing technical and/or analytical skills of the respective user role. The rule here is that with increasing self-service levels user roles with lower technical and/or analytical skills require greater support from the system. For example, in order to conduct a more-complex analytical task (e.g., harnessing new data sources) user roles that do not possess the required technical and/or analytical skills – such as data analysts – can benefit from a SSBIA system support. For instance, tools like Power BI (Microsoft, 2020a), KNIME, Alteryx Designer (Alteryx, 2020) or Tableau can enable data analysts to create mashups or access analytical functions on their own. Hereby, Power BI is a tool to easily create and share dashboards. In turn, Alteryx Designer is similar to KNIME and offers the possibility to create node based analytical workflows. In addition, by providing system support to casual user and data analysts, savvy user roles such as data scientists and/or IT departments would be relieved from helping those user roles. The freed capacities could be used for conducting other tasks. In contrast, as data scientists are typically capable of writing code for their analysis, they do not require support in this regard by a SSBIA system. Still, they can also benefit, for instance, by receiving access to already created reports or reusing existing statistical models with the help of model management warehouses in order to conduct their (complex) analytical tasks easier and faster than before (Chard et al., 2019; Schuff et al., 2018; Li et al., 2018). Power users can also benefit from the collection of user data. In particular, metadata of user data are collected providing a fingerprint of software usage and enabling a classification schema of how power user (i.e., data analysts) can be assisted in their daily work (Bilalli et al., 2016).

### 4.2.3 IT Users

IT users are in charge of providing the necessary infrastructure for working with data and thus act as enablers for the usage of such systems (De Mauro et al., 2018). Hereby, most studies emphasize the role of engineers and developers. In particular, the responsibility of engineers is varying from saving huge amounts of data in a database (e.g., Abelló et al., 2013) over providing data streams via pipelines for real-time data processing (e.g., Zehnder and Riemer, 2018) to management and governance of data (e.g., Bilalli et al., 2016). Developers, on the other hand, focus on the implementation of data-intensive systems in form of software or tools to support the data enriching and/or analysis process (Eckerson, 2012). These tasks include the design, development, or modification of systems that enable data analysts or data scientists to perform their tasks (Imhoff and White, 2011).

## 4.3 Experience

The majority of SSBIA initiatives target *novices* (67% of the studies). Still, it is important to manage the knowledge of expert users (33%) within SSBIA systems.

### 4.3.1 Novice Users

Novice users can reach acceptable results in analytical procedures and tasks with proper training (Lismont et al., 2019). In this line, education programs can implement a data-driven mindset throughout the organization (Lennerholt et al., 2018). Hereby, education programs have a bright scope ranging from explanations over dedicated training courses (e.g., Johannessen and Fuglseth, 2016) to business simulation games (e.g., Poonnawat and Lehmann, 2014).

### 4.3.2 Expert Users

In turn, Sulaiman and Gómez (2018) suggest using the knowledge of expert users to support novice users by providing recommendations to solve analytical tasks. Hereby, recommendations cover both unsupervised (e.g., clustering of unstructured data; Spahn et al., 2008) and supervised analytical tasks (e.g., performing decision trees in structured data; Zschech et al., 2019). The underlying rational of this approach resides in the idea of transferring expert knowledge represented by analysis paths for corresponding analytical tasks. Afterwards, these paths are provided as recommendations to novice or expert user. For example, to achieve a better data understanding, new data sources can be recommended if they were already applied for similar tasks in the past (Englmeier and Román, 2014). As a result, the correlation of often used and aggregated data sources can be exploited for the task at hand.

| | Perspective | | | User Role | | | | Experience | | Levels of Self-Service | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | Usage of information | | Creation of information | | Creation of information resources | | |
| | Artefact-Centric | User-Centric | Governance-Centric | Casual User | Data Analyst | Data Scientist | IT User | Novice | EXpert | Access to reports | Access to drill-anywhere reporting | Creation of reports | Access to analytical functions | Creation of mashups | Harnessing of new data sources | Deployment of analytical functions |
| Abelló et al., 2013 | X | X | | X | | | | X | | X | X | X | | X | X | |
| Ahmed et al., 2013 | X | X | | X | X | X | X | X | X | | | X | | X | X | |
| Alpar & Schulz, 2016 | | X | X | X | X | X | | X | X | X | X | X | X | X | X | |
| Arapakis et al. 2019 | | X | | X | X | | | X | | | | X | X | X | X | X |
| Badura & Schulz, 2018 | | X | | X | | | | | | | | | X | X | | |
| Bani-Hani et al., 2017 | | X | | X | X | | | X | X | X | | X | | X | | |
| Bani-Hani et al., 2018a | | X | | X | X | X | X | | | X | | X | X | X | X | |
| Bani-Hani et al., 2018b | | X | | X | X | | | X | X | X | | X | X | X | | |
| Behera & Swain 2019 | | X | | X | X | | | X | | | | | | | | |
| Behringer et al., 2018 | | X | X | | X | | | | | X | | X | X | X | X | |
| Berthold et al., 2010 | X | X | | X | X | | | X | | | | X | | X | X | |
| Bilalli et al., 2016 | X | X | | | X | | X | | | | | | | | | |
| Bilalli et al., 2018 | X | | | | X | | | | | | | | X | X | X | |
| Chard et al., 2019 | X | | | | | X | X | | X | | | | X | X | | X |
| Chouder et al., 2017 | X | | | X | X | X | | X | | | X | X | | X | X | |
| Chouder et al., 2019 | X | | | X | X | | | X | | | X | X | | X | X | |
| Clarke et al., 2016 | | X | X | X | X | X | X | X | X | X | X | X | | X | | |
| Convertino & Echenique, 2017 | | X | | | X | X | | X | X | | | X | | X | | |
| Daradkeh & Al-Dwairi, 2017 | | X | | X | X | | | | | | | | | | | |
| De Mauro et al. 2018 | | | | X | X | X | X | X | X | | | | | | | |
| Dinsmore 2016 | X | X | X | X | X | X | | X | X | X | | X | X | X | X | |
| Eckerson 2012 | | X | X | X | X | X | | X | X | X | X | X | X | X | X | |
| Eckerson 2009 | | X | X | X | X | X | | X | X | X | X | X | | X | X | |
| Englmeier & Román, 2014 | X | | X | X | | | | X | | X | | X | X | X | X | |
| Gröger, 2018 | X | X | X | X | X | X | | X | X | X | | X | X | X | X | X |
| Hilal et al., 2018 | X | | | | X | | | X | | | X | X | | X | X | |
| Horvath et al., 2014 | X | X | X | X | | | | X | | X | X | | | | X | |
| Ibáñez et al., 2017 | X | | | X | | | | | | | X | X | | X | X | |
| Imhoff & White, 2011 | | X | X | X | X | | X | X | | X | X | X | X | X | X | |
| Jie et al. 2018 | | X | | X | | | | | | | | X | | X | | |
| Johannessen & Fuglseth, 2016 | | X | | X | | | | X | | X | X | X | | X | X | |
| Johansson et al., 2015 | | X | | X | X | X | | X | X | X | | X | | X | X | |
| Kridel & Dolk, 2013 | X | X | | | X | X | | | | | | | X | X | | X |
| Lennerholt et al., 2018 | | X | X | X | X | X | | X | | X | | X | | X | | X |
| Li et al., 2017 | X | X | | X | X | | | X | | | | | X | X | | X |
| Lismont et al., 2019 | | X | | X | X | X | | X | X | | | | X | | | |
| Lizotte-Latendresse et al., 2018 | | X | | X | | | | | | X | X | X | X | X | | |
| Llave 2018 | | | | | X | X | | | | | | | | X | X | |
| Lousa et al., 2019 | | X | | X | X | | | X | X | X | | X | | X | X | |
| Marjanovic, 2015 | | X | | | | | | X | | X | | X | | | | |
| Mayer et al., 2015 | | X | | X | | | | | | X | X | X | X | | X | |
| Passlick et al., 2017 | X | X | X | X | X | X | | | | X | X | X | X | X | X | X |
| Poonnawat & Lehmann, 2014 | | X | | X | | | | X | | X | | X | | | | |
| Saeed & Abdinnour, 2011 | | X | | | | | | X | X | | | | | | | |
| Savinov, 2014 | X | X | | X | X | | | X | | | | | | X | X | |
| Schuff et al., 2018 | X | X | | X | X | | | X | | | | | X | | | |
| Smuts et al., 2015 | | X | | X | | | | X | | X | X | X | | X | X | |
| Spahn et al., 2008 | X | X | | X | | | | X | | X | | X | | X | X | X |
| Stanescu et al., 2015 | X | | | | X | X | X | | | | | X | X | X | X | X |
| Stodder, 2015 | | X | X | X | X | X | X | X | X | X | X | X | X | X | X | |
| Sulaiman & Gómez, 2018 | X | X | | X | X | | | X | X | X | | X | | X | | |
| Thornton & O'Flaherty, 2015 | | X | | X | X | | | X | | | | | X | | | |
| Varga et al., 2018 | X | X | | | | | | | | | | X | | X | X | |
| Vikram & Sherrymon, 2016 | X | | | X | | | | | | X | | | X | X | | |
| Gupta & Ashutosh, 2010 | | X | X | X | X | | | | | | | | | | X | |
| Vo et al., 2018 | X | X | X | X | | | | X | | X | | X | X | X | X | |
| Whitaker et al., 2018 | | X | | | X | | | X | | X | | X | | X | | |
| Zaghloul et al., 2014 | | X | | X | X | X | X | | | X | | X | X | X | X | X |
| Zehnder & Riemer, 2018 | X | X | | | | X | X | | X | | | | X | X | | X |
| Zscheech et al., 2019 | X | X | | X | X | | | X | X | | | | X | | | |
| **Total Number** | 27 | 49 | 15 | 46 | 41 | 22 | 11 | 40 | 20 | 31 | 17 | 40 | 28 | 46 | 34 | 11 |
| **Percentage** | 45% | 82% | 25% | 77% | 68% | 37% | 18% | 67% | 33% | 52% | 28% | 67% | 47% | 77% | 57% | 18% |

Table 1. Concept Matrix with Classification Results

## 4.4 Self-Service Levels

Due to the different user roles and corresponding diverging analytical tasks within SSBIA, the design of such systems requires a common understanding of the self-service levels. Following Alpar and Schulz (2016), the self-service levels include the following tasks: (1) *access to reports* (52% of the studies), (2) *access to drill anywhere reporting* (28%), (3) *creation of reports* (67%), (4) *access to analytical functions* (47%), (5) *creation of mashups* (77%), (6) *harnessing of new data sources* (57%), and (7) *deployment of analytical functions* (18%).

### 4.4.1 Usage of Information Resources

The **access to reports** requires the lowest level of self-reliance and system support from causal users (Alpar and Schulz, 2016). In particular, power users often create reports as a foundation for the underlying decision-making process of casual users. However, the dependency on reports increases if casual users are not self-reliant to conduct analytical tasks on their own. Thus, the "sharing of individual or collective insights becomes an important feature" because it could support the self-reliance of users and consumers alike (Marjanovic 2015, p. 4749). In this line, Passlick et al. (2017) propose an architecture that focuses on collaboration possibilities between heterogeneous user groups and a self-learning knowledge base to build a source for report recommendations. In addition, collaboration rooms are envisioned as a communication platform where casual user can make their achievements with their analytical tasks visible (Mayer et al., 2015).

The **access to drill "anywhere" reporting** is important for novices and could be subsumed by search, filter, and navigation facilities of SSBIA solutions (Eckerson, 2009; Smuts et al., 2015). For example, in tools like Power BI (Microsoft, 2020a) or Tableau (2019), users are able click on visualization areas like a bar or aggregated data points in a table to either filter the report on this dimension or expand the view to a data point level. However, SSBIA architectures must consider not only hierarchical drill "anywhere" capabilities (Lizotte-Latendresse and Beauregard, 2018) but also have to facilitate the problem investigations of casual users. In this direction, a critical amount of artifact-centric SSBIA studies provides technical solutions to apply the drill "anywhere" functionality on new data formats such as document stores (Chouder et al., 2017) and external data in triple form (Abelló et al., 2013). In addition, roll-up functionalities can complement drill "anywhere" reporting functionalities (Chouder et al., 2017). Although drill "anywhere" is an important functionality, it is rather underrepresented in SSBIA literature (only 28% of the studies).

### 4.4.2 Creation of Information

The first step for **creating a report** is to determine what can be achieved from a business perspective by defining objectives, business success criteria, and a corresponding project plan. On this basis, business requirements are collected from casual users and distributed to data analysts and data scientists stated in business-domain language (e.g., Daradkeh and Al-Dwairi, 2017). Moreover, studies discuss the relevance of data access and availability for the creation of information. For instance, Bani-Hani et al. (2017) propose SSBIA technologies to enable an equal access to organizational data for employees. In the second step, data are acquired and explored along their attributes and relationships (Berthold et al., 2010). Li et al. (2017) argue that SSBIA solutions are mainly supported for power and/or IT users leaving out business requirements of casual users. They tried to bridge this semantic gap by proposing an ontology-based approach to translate business requirements into data-related requirements. For example, predicting future sales performances depends on factors such as customers, suppliers, and/or marketing strategy. The mapped business-domain language is then represented by data which are acquired from various data sources, aggregated, and assigned to the underlying business problem.

The **access to analytical functions** refers to the application and testing of those on the previously created data sets in order to find the best fitting model and parameters (Dinsmore, 2016). Within the SSBIA literature, casual users should be empowered to apply analytical function because savvy power users are scares in the market (Chen et al., 2012) and organizations need to remain competitive (Bucher et al., 2009). Hereby, we could identify two approaches discussed in SSBIA literature: Firstly, model management warehouses are used to capture experiences with the application of algorithms to specific problem domains in form of trained models in order to bridge the gap between casual and power users (Chard et al., 2019; Schuff et al., 2018). As a prerequisite, an ontology-based data mining model management is necessary (Li et al., 2018). Secondly, solutions target casual users by offering non-code and drag-and-drop interfaces (Zehnder and Riemer, 2018). Hereby, models are built by chaining nodes visually which support the reuse of chains in other contexts. Such approaches also seem viable because scholars have shown that models trained by casual users are not necessarily worse than those trained by power users (Badura and Schulz, 2018; Lismont et al., 2019).

### 4.4.3 Creation of Information Resources

The **creation of mashups** represents a well-known data preparation task in the broader domain of BIA. Hereby, data are enriched and cleaned. In addition, inclusion or exclusion criteria are applied to gather the most useful data set (e.g., Abelló et al., 2013). Organizations typically provide full tool support for data preparation for experienced data scientists, but they only make up to 2% of the workforce within an organization (Eckerson, 2012). In turn, data analysts and inexperienced casual users are often not well-supported or have to deal with complex BIA tool landscapes. Against this backdrop, Savinov (2014) proposes a SSBIA tool for data integration because such operations are typically difficult for casual users. In particular, a unified data model is used, which enables casual users to create data mash-ups in a drag-and-drop manner by applying operations to existing elements. Moreover, Berthold et al. (2010) promote a global data model that is combined with ad-hoc and collaborative analytical capabilities throughout the organization. The global data model is responsible for translating between IT and business terminology. It refers to a mapping function to identify relevant key performance indicators and create user-specific context information that is directly linked to the underlying data. Technically, a flexible data integration methodology for the data infrastructure is provided to be capable of modifying data or linking it to business models by defined domain semantics. Furthermore, many vendors like Tableau (Tableau Prep; Tableau, 2020) and Alteryx (Alteryx Designer; Alteryx, 2020) start to offer easy-to-use data preparations tools. Microsoft has seamlessly integrated its data preparation tool Power Pivot (Microsoft, 2020b) into Excel and Power BI (Microsoft, 2020a).

In general, the **harnessing of new data sources** (for e.g. the combination of data from different departments) lead to richer insights for decision-making (Bani-Hani et al., 2017). SSBIA also enables users to explore new data in an autonomous way (Bani-Hani et al., 2018). New data can be in the form of structured data (e.g., from enterprise resource planning systems), unstructured data (e.g., from sensors accessible from manufacturing execution systems), or unstructured content (e.g., images from manufacturing quality systems; Gröger, 2018). However, the (automatic) incorporation of external data are likely to lead to meaningless queries and summarizability problems, especially with multidimensional data cubes (Ibánez et al., 2017). We identified a research stream that addresses this problem by proposing an open online analytical processing cube that considers data in the form of triplets (Abello et al., 2013; Hilal et al., 2018; Ibáñez et al., 2017). However, not only technical challenges occur: With the harnessing of new data sources in self-services, organizations and especially IT departments need to deal with data governance challenges (for details see Vo et al., 2018). For instance, business users want to cover decentral and locally managed data sources by business departments within their analytical procedures (Berthold et al., 2010).

Finally, the awareness for the possibility to **deploy analytical functions** by using self-services is rising (e.g., Kridel and Dolk, 2013; Lismont et al., 2019). The application of analytical functions results in statistical models. However, learned models can only create value once they are integrated into business processes. In this light, SSBIA architectures should consider model deployment. In particular, scholars suggest configurations that are exportable from the cloud to internal and local premises (Arapakis et al., 2019) or process-centric data analytics architecture with extended meta data support in order to simplify the control of deployment processes (Zaghloul et al., 2014). Zehnder and Riemer (2018) integrate ML algorithms into stream processing. Within the solution, more technical roles deploy analytical functions which enables domain experts to leverage ML in an easy-to-use graphical interface.

# 5 Future Research Avenues

In this section, an overview on promising research avenues for SSBIA is provided. Relying on our concept matrix (see Table 1), several research gaps can be identified. Following, we emphasize our insights and discuss potential solutions to address them.

(1) **Understand the Trade-Off between Top-Down and Bottom-Up SSBIA Capabilities.** Firstly, as users have diverse knowledge within analytical tasks, SSBIA tools need to address diverse capabilities ranging from data analysis with pre-defined sets of data (e.g., Schuff et al., 2018) to data integration, preparation, and modeling (e.g., Abelló et al., 2013; Savinov, 2014). The technical infrastructure should enable all user groups to conduct their analytical tasks based on the underlying data. However, a typical "one size fits all" approach fails as casual users may be overcharged with the BIA capabilities, whereas power users may require a higher degree of flexibility or even more complex functions (Eckerson, 2012). Although a variety of user roles is addressed by SSBIA literature in our SLR (77 % casual user, 68% data analysts, 37 % data scientists, 18% IT users), future work still needs to understand the trade-off between top-down (less complex, more-rigid) and bottom-up (complex, more-flexible) SSBIA capabilities. Hereby, our SLR results have shown that SSBIA is not only important for enabling casual users to prepare and analyse data with easy-to-use BIA systems without the need for expert support (e.g., Lennerholt et al., 2018) but is also relevant to support power users to perform their (complex) analytical tasks easier and faster than before (e.g., Schuff et al., 2018). On these grounds, it seems important to extract the involved user roles with their corresponding skill levels, activities conducted, and tool classes used in order to setup the basis for an effective SSBIA system support. A promising research avenue can therefore be to extract and cluster BIA-related job postings (e.g., De Mauro et al., 2018) combined with interviews with real-world BIA experts and users in the field in order to shed light on this trade-off.

(2) **Define a Method for Implementing SSBIA within Organizations.** Secondly, there exists no method support for using and implementing SSBIA within organizations. Although the concept of the self-service levels by Alpar and Schulz (2016) sheds light on how self-reliant users can perform their tasks and what degree of system support is required, SSBIA research still lacks knowledge regarding the temporal sequence of tasks within the analytical investigation process. For example, the creation of a report is a necessary prerequisite for accessing it. However, oftentimes such task interdependencies lie hidden and are complex due to the involvement of various user roles demanding a careful assessment. Other well-established methods such as CRISP-DM (Chapman et al., 2000; Shearer, 2000) do also not provide full support for using and implementing SSBIA within organization as they appear to be too limited. For instance, CRISP-DM was particularly developed to create statistical models with data mining. From a tool perspective, software vendors such as KNIME already offer functionalities along the analytical investigation process. But such tools do also not offer method support (e.g., for mapping a domain-specific problem description to the class of data mining methods; Zschech et al., 2019). Finally, with the expansion of SSBIA in organizations, it gets more and more blurry which user roles conduct which tasks. Most SSBIA studies, however, do not describe well how collaboration or information sharing can be implemented (Bani-Hani et al., 2017). Against this backdrop, we suggest defining a method for cross-disciplinary teams of casual, power, and IT users, as well as to consider existing and/or new methods to deal with diverse data, tool landscapes, and reciprocal iterations between the user roles.

(3) **Investigate Success Factors for Data Management and Data Governance.** Thirdly, data management and data governance play a critical role in establishing SSBIA initiatives (e.g., Clarke et al., 2016; Eckerson, 2012; Gröger, 2018; Passlick et al., 2017). However, despite this relevance, only 25% of studies address this topic. Studies illustrate that if organizations use data according to well-defined procedures for data management and data governance, users may comprehend the value of the underlying data reducing the probability of creating a "shadow IT" that is problematic with regard to having access to or relying on the right data (e.g., Lennerholt et al., 2018; Stodder, 2015). On these grounds, future research can try to understand the success factors of implementing data management and governance procedures within SSBIA initiatives, especially with regard to

how easy-to-use data access can be established and how an appropriate data quality can be implemented for analytical tasks. As already mentioned in the results section, in data lakes, casual user may struggle with data quality issues because data are typically stored schema-free and in a raw format. To enable easy and quality-proven data access to data lakes for casual users, data needs to be modelled. For instance, the data vault modelling technique was promising for a broad variety of use case implementations. However, the data vault reference lacks guidelines and best practices for the application as enterprise-wide modelling technique for data lakes (Giebler et al. 2019). These circumstances can be used as an interesting starting point for further investigations.

(4) **Support Casual Users' Self-Reliance within the Analytical Investigation Process.** While there is a great deal of past and current SSBIA tools, casual users still need support to consume insights or to conduct analytical tasks on their own. Furthermore, the access to analytical functions is (still) not well supported (Schuff et al., 2018). First attempts propose to rely on model management warehouses (Chard et al., 2019; Schuff et al., 2018) in order to enable the reuse of existing models trained by power users. Future work can therefore implement SSBIA guidance services by relying on domain-knowledge from power users. As suggested by Sulaiman and Gómez (2018), analysis paths from power users can be logged (e.g., in a data lake) and leveraged as a foundation for recommendations. For instance, it can be showcased what kind of data, algorithms, and performance criteria is used by power users for similar tasks. In the long run, such recommendation support can improve casual users' self-reliance and knowledge of diverse analytical tasks within the underlying investigation process (e.g., Lismont et al., 2019).

(5) **Develop Effective Multi-Sensory User Interfaces for Immersive Collaboration Environments.** Finally, we see the need for a better collaboration between casual and power users due to the semantic gap between both user groups. Both groups are not "fluent" in the domain language of each other leading to obstacles to meet the defined performance criteria (Lennerholt et al., 2018). The large body of knowledge about immersive analytics can be made accessible as current SSBIA literature (as illustrated in the SLR) neglected the appropriate consideration of this stream. Immersive analytics is an emerging academic stream which analyses the effects of new interaction and display technologies on analytical reasoning and decision-making processes (Chandler et al., 2015). Technologies in form of touch displays, virtual and augmented reality devices, and other sensors are used as a backbone to develop multi-sensory interfaces for collaboration enabling diverse user groups to immerse themselves in their data (Chandler et al., 2015). For instance, users may talk to BIA dashboards through digital assistants. Such a platform can be made accessible for SSBIA and used to collaborate more effectively on a cross-disciplinary basis. Hereby, data analysts can improve their data models by collaborating with casual users to identify relevant features and/or supplement novel data sources based on the suggestions of data scientists.

# 6 Conclusion

Our SLR was structured along three contributions: First, categories for the concept matrix were defined to structure multiple SSBIA characteristics. Second, we provided a state-of-the art overview by typecasting existing studies within SSBIA literature. Third, we offered recommendations by formulating avenues for future research to be well-directed. We are aware that our paper has some limitations. Any bias in the selection of the search string might result in a bias of the reviewed studies. To reduce this probability, our SLR and the subsequent search process are based on methodological recommendations prescribed in the literature (i.e., Kitchenham and Charters, 2007; Webster and Watson, 2002). All choices during the plan, conduct, and report stage are made explicit. As next steps, we plan to conduct interviews with BIA experts and users to validate and extend our SLR outcomes and supplement the results with contemporary issues from practitioners. Although some authors rely on theories and concepts (e.g., Jie et al., 2018; Saeed and Abdinnour, 2011) or illustrate the SSBIA business value, the domain lacks a solid foundation with regard to systematic methods and an established theory still needs to materialize to date. We hope that our SLR can serve as a reference for scholars in the broader field of SSBIA and the aspects that should be considered when investigating related concepts and solutions.

## References


Abelló, A., Darmont, J., Etcheverry, L., Golfarelli, M., Mazón, J.-N., Naumann, F., Pedersen, T., Rizzi, S. B., Trujillo, J., Vassiliadis, P., and Vossen, G. (2013). "Fusion Cubes: Towards Self-Service Business Intelligence," *International Journal of Data Warehousing and Mining,* (9:2), pp. 66–88. (DOI: 10.4018/jdwm.2013040104).

Ahmed, S. and Ruhi, U. (2013). "Towards a functional taxonomy of enterprise business intelligence mashups" In: *2013 Second International Conference on Informatics & Applications (ICIA):* pp. 98–103.

Alpar, P., and Schulz, M. (2016). "Self-Service Business Intelligence," *Business and Information Systems Engineering (58:2),* pp. 151–155. (DOI: 10.1007/s1259901604246).

Alteryx (2020). *Alteryx Designer: Repeatable workflows for self-service data analytics.* URL:https://www.alteryx.com/products/alteryx-platform/alteryx-designer (visit on 29/03/2020).

Arapakis, I., Becerra, Y., Boehm, O., Bravos, G., Chatzigiannakis, V., Cugnasco, C., Demetriou, G., Eleftheriou, I., Etienne Mascolo, J., Fodor, L., Ioannidis, S., Jakovetic, D., Kallipolitis, L., Kavakli, E., Kopanaki, D., Kourtellis, N., Maawad Marcos, M., Martin de Pozuelo, R., Milosevic, N., Morandi, G., Pages Montanera, E., Ristow, G., Sakellariou, R., Sirvent, R., Skrbic, S., Spais, I., Vasiliadis, G. and Vinov, M. (2019). "Towards Specification of a Software Architecture for Cross-Sectoral Big Data Applications" In: *IEEE services 2019: Proceedings*, pp. 394–395.

Badura, D., and Schulz, M. (2018). "Datenkompetenzen für die Massen - Muss Self-Service-Data-Mining scheitern? ", *Conference: Lernen, Wissen, Daten, Analysen 2018 (LWDA).*

Bandara, W., Furtmueller, E., Gorbacheva, E., Miskon, S., and Beekhuyzen, J. (2015). "Achieving rigor in literature reviews: Insights from qualitative data analysis and tool-support," *Communications of the Association for Information Systems,* (37:1), pp. 154–204.

Bani-Hani, I., Pareigis, J., Tona, O., and Carlsson, S. (2018a). "A holistic view of value generation process in a SSBI environment: a service dominant logic perspective," *Journal of Decision Systems* (27:1), pp. 46–55. (DOI: 10.1080/12460125.2018.1468155).

Bani-Hani, I., Tona, O., and Carlsson, S. (2018b). "From an information consumer to an information author: a new approach to business intelligence," *Journal of Organizational Computing and Electronic Commerce,* (28:2), pp. 157–171. (DOI: 10.1080/10919392.2018.1444358).

Bani-Hani, I., Deniz, S., and Carlsson, S. (2017). "Enabling Organizational Agility Through Self-Service Business Intelligence: The Case of a Digital Marketplace," PACIS 2017 Proceedings (148).

BARC (2019). "BI Trend Monitor - Business Application Research Center," URL: http://barc-research.com/research/bi-trend-monitor/ (visit on 04/10/2019).

Behera, R.K. and Swain, A.K. (2019). "Big data real-time storytelling with self-service visualization" In: *Advances in Intelligent Systems and Computing,* (813), pp. 405-415.

Behringer, M., Hirmer, P., and Mitschang, B. (2018). "A Human-Centered Approach for Interactive Data Processing and Analytics," *Enterprise Information Systems,* (19), pp. 498–514. (DOI: 10.1007/978331993375723).

Berthold, H., Rösch, P., Zöller, S., Wortmann, F., Carenini, A., Campbell, S., Bisson, P., and Strohmaier, F. (2010). "An architecture for ad-hoc and collaborative business intelligence," In: EDBT'10 *Proceedings of the 2010 EDBT/ICDT Workshops on Data Semantics – DataSem,* (13), pp. 1-6. (DOI: 10.1145/1754239.1754254).

Bilalli, B., Abelló, A., Aluja-Banet, T., Munir, R.F. and Wrembel, R. (2018). "PRESISTANT: Data Pre-processing Assistant", In: *Information systems in the big data era: CAiSE Forum 2018*, Lecture Notes in Business Information Processing, (317), pp. 57–65.

Bilalli, B., Abelló, A., Aluja-Banet, T., and Wrembel, R. (2016). "Towards Intelligent Data Analysis: The Metadata Challenge," In: *Proceedings of the International Conference on Internet of Things and Big Data*, pp. 331–338. (DOI: 10.5220/0005876203310338).

Bondarouk, T., Parry, E. and Furtmueller, E. (2017). "Electronic HRM: four decades of research on adoption and consequences," *The International Journal of Human Resource Management*, (28:1), pp. 98-131. (DOI: 10.1080/09585192.2016.1245672).

Bose, R. (2009). “Advanced analytics: Opportunities and challenges”, *Industrial Management and Data Systems* (109:2), pp. 155-172. (DOI: 155-172. 10.1108/02635570910930073).

Bucher, T., Gericke, A., and Sigg, S. (2009). “Process-centric business intelligence,” *Business Process Management Journal,* (15:3), pp. 408-429.

Chandler, T., Cordeil, M., Czauderna, T., Dwyer, T., Glowacki, J., Goncu, C., Klapperstueck, M., Klein, K., Schreiber, F., Wilson, Elliot, (2015). “Immersive Analytics,” *Big Data Visual Analytics (BDVA*), IEEE, pp. 73-80. (DOI: 10.1109/BDVA.2015.7314296).

Chapman, P., Clinton, J. M., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. R. H., Wirth, R. (2000). *CRISP-DM 1.0: Step-by-step data mining guide.*

Chard, R., Li, Z., Chard, K., Ward, L., Babuji, Y., Woodard, A., Tuecke, S., Blaiszik, B., Franklin, M. and Foster, I. (2019). “DLHub: Model and Data Serving for Science”, In: *IPDPS 2019: Proceedings*, pp. 283–292.

Chen, H., Chiang, R. H. L., and Storey, V. C. (2012). “Business intelligence and analytics: from big data to big impact,” *MIS quarterly,* (36:4), pp. 1165–1188. (DOI: 10.2307/41703503).

Chouder, M. L., Rizzi, S., and Chalal, R. (2017). “Enabling self-service BI on document stores,” In: *CEUR Workshop Proceedings in EDBT/ICDT Joint Conference,* (1810).

Clarke, P., Tyrrell, G., and Nagle, T. (2016). “Governing self-service analytics,” *Journal of Decision Systems* (25:1), pp. 145–159. (DOI: 10.1080/12460125.2016.1187385).

Convertino, G., and Echenique, A. (2017). “Self-Service Data Preparation and Analysis by Business Users: New Needs, Skills, and Tools,” *ACM CHI EA - Conference Extended Abstracts on Human Factors in Computing Syste*ms, pp. 1075-1083. (DOI: 10.1145/3027063.3053359).

Daradkeh, M., and Al-Dwairi, R. M. (2017). “Self-Service Business Intelligence Adoption in Business Enterprises,” *International Journal of Enterprise Information Systems,* (13:3), pp. 65–85. (DOI: 10.4018/IJEIS.2017070105).

De Mauro, A., Greco, M., Grimaldi, M., and Ritala, P. (2018). “Human resources for Big Data professions: A systematic classification of job roles and required skill sets,” *Information Processing & Management* (54:5), pp. 807–817. (DOI: 10.1016/j.ipm.2017.05.004).

Dinsmore, T. W. (2016). “Disruptive Analytics: Charting your strategy for next-generation business analytics,” Berkeley, CA: Apress.

Eckerson, W. (2012). *Business-driven BI: Using new technologies to foster self-service access to insights.* BeyeRESEARCH, pp. 1–30.

Eckerson, W. (2009). Self-service BI. TDWI Checklist Report, TDWI Research

Englmeier, K., and Román, H. (2014). “Self-Service Data Discovery and Information Sharing - Fostering the Engineering Capacity of the Data-Driven Mindset,” In: *Proceedings of the International Conference on Knowledge Management and Information Sharing,* pp. 339–345. (DOI: 10.5220/0005153303390345).

Giebler, C., Gröger, C., Hoos, E., Schwarz, H., and Mitschang, B. (2019). “Modeling Data Lakes with Data Vault: Practical Experiences, Assessment, and Lessons Learned”, In: *Conceptual Modeling,* , pp. 63–77.

Gröger, C., and Hoos, E. (2019). “Ganzheitliches Metadatenmanagement im Data Lake: Anforderungen, IT-Werkzeuge und Herausforderungen in der Praxis,” In: *Proceedings 18. Fachtagung Datenbanksysteme für Business, Technologien and Web (BTW),* Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, pp. 435–452.

Gröger, C. (2018). “Building an Industry 4.0 Analytics Platform; Practical Challenges, Approaches and Future Research,” *Datenbank-Spektrum,* (18:1), pp. 5–14. (DOI: 10.1007/s1322201802731).

Gupta, V., Saxena, A. (2010). “Personalized Data Set For Analysis”, In*: IJDMS (2:4)*, pp. 37–44. DOI: 10.5121/ijdms.2010.2404.

Hamari, J., Koivisto, J., and Sarsa, H. (2014). “Does gamification work? - A literature review of empirical studies on gamification,” In: *Proceedings of 47th the Hawaii International Conference on System Sciences (HICSS)*, pp. 3025–3034.

Hilal, M., Schuetz, C.G., and Schrefl, M. (2018). “Using superimposed multidimensional schemas and OLAP patterns for RDF data analysis,” *Open Computer Science*, (8:1), pp. 18–37.

Horvath, M. M., Rusincovitch, S. A., Brinson, S. Shang, H. C., Evans, S., and Ferranti, J. M. (2014). “Modular design, application architecture, and usage of a self-service model for enterprise data delivery: The Duke Enterprise Data Unified Content Explorer (DEDUCE),” *Journal of Biomedical Informatics. Special Section: Methods in Clinical Research Informatics,* (52), pp. 231–242. (DOI: 10.1016/j.jbi.2014.07.006).

Ibáñez, L.-D., Mazón, J.-N., and Simperl, E. (2017). “Towards Semantic Assessment of Summarizability in Self-service Business Intelligence,” *European Conference on Advances in Databases and Information Systems - New Trends in Databases and Information Systems,* (767:1), pp. 174–185. (DOI: 10.1007/978331967162818).

Imhoff, C., and White, C. (2011). Self-service Business Intelligence: Empowering Users to Generate Insights, *TDWI Best practices report*.

Jie, L.Y., Wong, D.H.-T., Zain, Z.M., Sjarif, N.N.A., Ibrahim, R., and Maarop, N. (2018). “Metrics and benchmarks for empirical and comprehension focused visualization research in the sales domain”, *Indonesian Journal of Electrical Engineering and Computer Science,* (12:3), pp. 1340–1348.

Johannessen, T. V., and Fuglseth, A. M. (2016). “Challenges of Self-Service Business Intelligence,” NOKOBIT (24:1), pp. 151–155. (DOI: 10.1007/s1259901604246).

Johansson, B., Alkan, D., and Carlsson, R. (2015). “Self-service BI does it change the rule of the game for BI systems designers,” *ILOG’15: 8th Workshop on Information Logistics and Knowledge Supply,* (1420), pp. 48–61.

KDNuggets (2014). *What main methodology are you using for your analytics, data mining, or data science projects?* URL: https://www.kdnuggets.com/polls/2014/analytics-data-mining-data-science-methodology.html (visit on 04/20/2019).

Kitchenham, B., and Charters, S. (2007). “Guidelines for performing Systematic Literature Reviews in software engineering,” *EBSE Technical Report*. EBSE-2007-01.

KNIME (2019). *Open for Innovation*. URL: https://www.knime.com/ (visit on 10/04/2019).

Kridel, D., and Dolk, D. (2013). “Automated self-service modeling: Predictive analytics as a service,” *Journal of Information Systems and e-Business Management,* (11:1), pp. 119–140. (DOI:10.1007/s10257-011-0185-1).

Lennerholt, C., Van Laere,J., and Söderström, E. (2018). “Implementation Challenges of Self Service Business Intelligence: A Literature Review,” In: *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS)*, pp. 5055–5063.

Li, Y., Thomas, M., A., and Osei-Bryson, K.-M. (2017). “Ontology-based data mining model management for self-service knowledge discovery,” *Journal of Research and Innovation - Information Systems Frontiers,* (19:4), pp. 925–943. (DOI: 10.1007/s10796- 016- 9637- y).

Lismont, J., Van Calster, T., and Óskarsdóttir, M. (2019). “Closing the Gap Between Experts and Novices Using Analytics-as-a-Service: An Experimental Study,” *Business Information Systems Engineering,* (6), pp. 1-15. (DOI: 10.1007/s12599-018-0539-z).

Lizotte-Latendresse, S., and Beauregard, Y. (2018). “Implementing self-service business analytics supporting lean manufacturing: A state-of-the-art review,” *IFAC-PapersOnLine,* (51:11), pp. 1143–1148. (DOI: 10.1016/j.ifacol.2018.08.436).

Llave, M.R. (2018). “Data lakes in business intelligence: Reporting from the trenches”, *Procedia Computer Science*, pp. 516.

Logi (2017). *Business Intelligence Adoption Report.* URL: https://www.logianalytics.com/report/2017-state-analytics-adoption/ (visit on 04/20/2019).

Lousa, A., Pedrosa, I., and Bernardino, J. (2019). “Evaluation and Analysis of Business Intelligence Data Visualization Tools” In: *14th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1–6.

Marjanovic, O. (2015). “From analytics-as-a-service to analytics-as-a-consumer-service: Exploring a new direction in business intelligence and analytics research,” *in Proceedings of the 48th Hawaii International Conference on System Sciences (HICSS),* pp. 4742–4751. (DOI: 10. 1109/HICSS.2015.564).

Mayer, J. H., Hartwig, J., Röder, A., and Quick, R. (2015). "Self-Service Management Support Systems: Findings from a New-Generation Manager Perspective," In: *Business Intelligence for New-Generation Managers,* pp. 113–136. (DOI: 10.1007/978-3-319-15696-5_9).

Miloslavskaya, N. and Tolstoy, A. (2016). "Big Data, Fast Data and Data Lake Concepts," *In Procedia Computer Science*, (88), pp. 300–305. DOI: 10.1016/j.procs.2016.07.439.

Microsoft (2020a). *Turn data into opportunity - Drive better business decisions across your organization with Microsoft Power BI*. URL: https://powerbi.microsoft.com/en-us/ (visit on 29/03/2020).

Microsoft (2020b). *Power Pivot: Powerful data analysis and data modeling in Excel.* URL: https://support.office.com/en-us/article/power-pivot-powerful-data-analysis-and-data-modeling-in-excel-a9c2c6e2-cc49-4976-a7d7-40896795d045?omkt=en-001&ui=en-US&rs=en-001&ad=US (visit on 29/03/2020).

Mitchell, T. (1997). Machine learning. McGraw-Hill Series in Computer Science.

Passlick, J., Lebek, B., and Breitner, M. H. (2017). "A Self-Service Supporting Business Intelligence and Big Data Analytics Architecture," In: *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI),* (80:9), pp. 1126–1140. (DOI: 10.5120/13893-1840).

Poonnawat, W., and Lehmann, P. (2014). "Using self-service business intelligence for learning decision making with business simulation games," In: *Proceedings of the 6th International Conference on Computer Supported Education (CSEDU),* (3), pp. 235–240. (DOI: 10.5220/0004941202350240).

Saeed, K. A., and Abdinnour, S. (2011). "Understanding post-adoption IS usage stages: An empirical assessment of self-service information systems," *Information Systems Journal*, (23:3), pp. 219–244. (DOI: 10.1111/j.1365-2575.2011.00389.x).

Savinov, A. (2014). "ConceptMix: Self-Service Analytical Data Integration Based on the Concept-Oriented Model," In: *Proceedings of the 3rd International Conference on Data Technologies and Applications (DATA 2014)*, pp. 78-84. (DOI: 10.5220/0005103700780084).

Schuff, D., Corral, K., Louis, R., and Schymik, G. (2018). "Enabling self-service BI: A methodology and a case study for a model management warehouse," *Information Systems Frontiers (20:2),* pp. 275–288. (DOI: 10.1007/s10796-016-9722-2).

Shearer, C. (2000). "The CRISP-DM model: the new blueprint for data mining," *Journal of data warehousing,* (5:4), pp. 13–22.

Spahn, M., Kleb, J., Grimm, S., and Scheidl, S. (2008). "Supporting business intelligence by providing ontology-based end-user information self-service," In: *Proceedings of the first international workshop on Ontology-supported business intelligence – (OBI),* pp. 1–12. (DOI: 10.1145/1452567.1452577).

Smuts, M., Scholtz, B., and Calitz, A. (2015). "Design Guidelines for Business Intelligence Tools for Novice Users", *SAICSIT 2015: Proceedings of the 2015 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists Knowledge through Technology,* pp. 1–15.

Stanescu, I. A., Stefan, V., Neagu, G., and Cîrnu, C. E. (2015). "Renewable energy decision support systems: The challenge of data integration," *Studies in Informatics and Control (24:2)*, pp. 191–200. (DOI: 10.24846/v24i2y201507).

Stodder, D. (2015). "Visual analytics for making smarter decisions faster–applying self-service business intelligence technologies to data-driven objectives," *TDWI Best Practices Report.*

Sulaiman, S., and Gómez, J. M. (2018). "Recommendation-based Business Intelligence Architecture to Empower Self Service Business Users," *Multikonferenz Wirtschaftsinformatik*, pp. 1–12.

Tableau (2019). "Tableau: Business Intelligence and Analytics Software," URL: https://www.tableau.com/ (visit on 04/10/2019).

Thornton, C., and O'Flaherty, B. (2015). "Improving Customer Centric Design for Self-Service Predictive Analytics," *International Conference on Design Science Research in Information Systems: New Horizons in Design Science: Broadening the Research Agenda (DESRIST),* pp. 230-245. (DOI: 10.1007/978-3-319-18714-3_15).

Trigo, A., Belfo, F., and Estébanez, R. P. (2014). "Accounting Information Systems: The Challenge of the Real-time Reporting," In: *Procedia Technology,* (16:1), pp.118–127. (DOI: 10.1016/j.protcy.2014.10.075).

Varga, J., Romero, O., Pedersen, T. B., and Thomsen, C. (2018). "Analytical metadata modeling for next generation BI systems," *Journal of Systems and Software,* (144:1), pp. 240–254. (DOI: 10.1016/j.jss.2018.06.039).

Srivatsa, Vikram N.; Mathew, Sherrymon (2016). SMART Analytical Platform. In: *IUP Journal of Computer Sciences*, (10:1-2), p. 43-51.

Vo, Q.D., Thomas, J., Cho, S., De, P., and Choi, B.J. (2018). "Next Generation Business Intelligence and Analytics", In: *ICBIM '18 Proceedings of the 2nd International Conference on Business and Information Management,* pp. 163-168.

Webster, J., and Watson, R. T. (2002). "Analyzing the past to prepare for the future: Writing a literature review," *MIS quarterly*, (26:2), pp. xiii–xxiii.

Watson, H. J. (2009). "Tutorial: Business Intelligence – Past, Present, and Future," *Communications of the Association for Information Systems* (25:1), pp. 488-510, (DOI: 10.17705/1CAIS.02539).

Whitaker, T. J. Mayo, C. S., Haddock, M. G., Miller, R. C., Corbin, K. S., Neben-Wittich, M., Leenstra, J. L., Laack, N. N., Fatyga, M., Schild, S. E., Vargas, C. E., Tzou, K. S., Hadley, A. R., Burskirk, S. J., and Foote, R. L. (2018). "Data collection of patient outcomes: One institution's experience," *Journal of Radiation Research,* (59:1), pp. i19–i24. (DOI: 10.1093/jrr/rry013).

Zaghloul, M. M., Ali-Eldin, A., and Salem, M. (2014). "A process-centric data analytics architecture," *9th International Conference on Informatics and Systems (INFOS),* pp. 34-39, (DOI: 10.1109/INFOS.2014.7036705).

Zehnder, P., and Riemer, D. (2018). "Modeling self-service machine-learning agents for distributed stream processing" *IEEE International Conference on Big Data (Big Data)*, pp. 2203–2212. (DOI: 10.1109/BigData.2017.8258170).

Zschech, P., Heinich, K., Horn, R. and Höschele, D. (2019). "Towards a Text-based Recommender System for Data Mining Method Selection", 25th *American Conference on Information Systems* (AMCIS).