5-11-2023

Guidance in Business Intelligence & Analytics Systems: A Review and Research Agenda

Jonas Gunklach
Karlsruhe Institute of Technology (KIT), jonas.gunklach@kit.edu

Mario Nadj
TU Dortmund University, mario.nadj@tu-dortmund.de

GUIDANCE IN BUSINESS INTELLIGENCE AND ANALYTICS SYSTEMS: A REVIEW AND RESEARCH AGENDA

Research Paper

Gunklach, Jonas, Human-Centered Systems Lab, Karlsruhe Institute of Technology, Karlsruhe, Germany, jonas.gunklach@kit.edu

Nadj, Mario, Business & Information Systems Engineering, TU Dortmund University, Dortmund, Germany, mario.nadj@tu-dortmund.de

Abstract

While the data amount grows exponentially, the number of people with analytical and technical skills is only slowly increasing. This skill gap is putting pressure on the labor market and increasing the need for personnel with these skills. At the same time, companies are forced to think of alternative ways to empower their less-skilled workforce to take on Business Intelligence and Analytics (BI&A) tasks. One promising attempt to address these challenges may turn to the concept of guidance. However, the current body of research on guidance in BI&A systems is scattered and lacks a structured investigation from which future research avenues can be derived. To address this gap, this article analyzes five categories, namely BI&A phases, guidance degree, guidance generation, user roles, and interactivity form. Reviewing 82 articles, our contribution is to synopsize articles on guidance in BI&A systems and to suggest five research avenues.

Keywords: Guidance, Business Intelligence, Analytics, Literature Review

1 Introduction

Today's organizations rely on Business Intelligence & Analytics (BI&A) to make smarter and faster decisions (Trieu et al., 2022). BI&A refers to all "techniques, technologies, systems, practices, methods, and applications that analyze critical business and market and make timely business decisions" (Chen et al., 2012, p. 1166). BI&A helps companies make informed decisions by offering information about historic, current, and future performance (e.g., in the form of future trends, expected demands, or customer behavior) (Rikhardsson and Yigitbasioglu, 2018). As a result, the market size for BI&A software applications is forecast to grow to more than 18 billion in 2025 (Vailshery, 2022). But to make use of large data amounts and hence support decision-making, data must be analyzed, and this requires analytical and technical skills on the part of humans (Lennerholt et al., 2021). Commonly, skilled users collect and analyze the data to create reports on which less-skilled users can base their decisions on (Imhoff and White, 2011). However, while the available data is growing exponentially simultaneously, the number of people with analytical and technical skills is increasing only slowly (Awasthi and George, 2020). Thus, this skill gap puts pressure on the labor market and increases the need for human resources with these skills (Agrawal et al., 2020) while also forcing companies to think of alternative ways to empower their less-skilled workforce.

In particular, organizations have already begun to equip less-skilled users with self-service tools to enable them in preparing and running their own analyses (Alpar and M. Schulz, 2016). Prominent examples refer to tools like KNIME or Tableau (Michalczyk et al., 2020). KNIME supports interactive data analysis and enables the inclusion of various machine learning (ML) modules and methods, meanwhile Tableau's software, for instance, helps users create visualizations using drag-and-drop (Michalczyk et al., 2020). However, recent research has highlighted several user-related challenges associated with self-service capabilities, ranging from problems accessing and using the data to difficulties using the current tools to even issues in interpreting the resulting outcomes due to users' limited skills (Lennerholt et al., 2021). One promising attempt to address these challenges could relate to the concept of guidance (Ceneda et al., 2019). Guidance is a computer-assisted process that aims to actively resolve a skill gap encountered by users (Ceneda et al., 2017). Hereby, guidance can assist users in "decision making, problem-solving, and task execution during system use by providing suggestions and information" (Morana et al., 2017, p. 31). In the past, several articles have examined the effects of deliberately including guidance features into information systems and concluded that guidance can have positive effects, for instance, on decision accuracy and speed (Shen et al., 2012). Research also recognized the need to consider BI&A systems that offer guidance to support users in their underlying tasks (Michalczyk et al., 2020). For instance, Ceneda et al. (2017) emphasized the need to guide users in the visual analysis process. Thus, BI&A systems equipped with guidance capabilities can be found throughout the literature (e.g. Shi et al., 2021; Zschech et al., 2020). However, no systematic literature review (SLR) is currently available that offers an overview of such BI&A systems. Further, research is scattered and a structure could help to suggest avenues for future research. On this basis, we present a state-of-the-art overview and identify future research avenues following well-established methodological guidelines (e.g. Kitchenham and Charters, 2007; Webster and Watson, 2002). Specifically, our analysis investigates the (1) BI&A phases these systems support, characteristics that describe the guidance concept (i.e., (2) guidance degree and (3) guidance generation), (4) targeted user roles, as well as (5) the interactivity form.

In this article, we analyze 82 articles that cover the design, implementation, or evaluation of guidance in BI&A systems. We formulated the following research question: *What is the state-of-the-art of guidance in Business Intelligence & Analytics systems, and what are potential avenues 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 SLR. We then present the results of our study in section four. In section five, we outline future research avenues before concluding our article in section six.

2 Foundations

BI&A "encompasses a variety of technologies and methodologies that enable organizations to collect data from internal and external sources, prepare it for analysis, develop and run queries against the data, and create reports, dashboards, and data visualizations to make the results available to end-users" (Rikhardsson and Yigitbasioglu, 2018, p. 3). Hereby, BI&A relies on various data collection, extraction, and analysis technologies (Lim et al., 2013). Following literature (e.g. Chae and Olson, 2013; Chaudhuri et al., 2011; Miloslavskaya and Tolstoy, 2016), there are four essential technological elements of BI&A in organizations: (1) infrastructure (e.g., data warehouses or data lakes), (2) data management (e.g., integration of internal and external data), (3) analytical tools (e.g., support statistical analysis of data), and (4) reporting tools (e.g., provide information to decision-makers). The fundamental purpose of these solutions is to gather, process, and analyze data in order to improve decision-making (Chen et al., 2012). An essential factor in the successful implementation of BI&A solutions is whether a system is capable of supporting the skills of the targeted user roles (Lismont et al., 2019). Organizations typically face the challenge of addressing different user roles with diverse skill levels for various analytical demands (Alpar and M. Schulz, 2016). According to Michalczyk et al. (2020), there exists a spectrum of job roles involved in different BI&A phases, each with a different combination of business, analytical, and technical skills. For that, organizations have already begun to equip less-skilled users with self-service tools to enable them in preparing and running their own analysis (Alpar and M. Schulz, 2016). However, recent research has highlighted several user-related challenges associated with self-service capabilities, ranging from problems in accessing and using the data (e.g., using correct data queries and making data sources easy to access) to difficulties using the current tools to even issues in interpreting the resulting outcomes due to users' limited skills (Lennerholt et al., 2021).

This is where guidance could make a difference. Specifically, the use of guidance can have a positive impact on user system acceptance (Ye and Johnson, 1995), user satisfaction (Huguenard and Frolick, 2001), or effectiveness and efficiency of decision-making (Gregor and Benbasat, 1999; Shen et al., 2012; Silver, 1991). In general, guidance refers to "the act or process of guiding" (Merriam-Webster, 2022) with the aim to address a problem (Ceneda et al., 2017; Lexico, 2022). Moreover, research highlighted the supportive nature of guidance in different contexts by means of interactivity (Smith and Mosier, 1986), dynamic support (e.g., for data exploration), or suggestions of appropriate resources (e.g., human experts or infrastructure) to conduct a respective task (H.-J. Schulz et al., 2013). Following Ceneda et al. (2017), guidance refers to as a concept with the goal of providing support to users (e.g., when exploring data or when finding the best visualizations for presenting data). Hereby, three guidance degrees are distinguished: (1) orienting, (2) directing, and (3) prescribing. For instance, Tableau provides guidance to users in form of several features. Specifically, imagine a user would like to visualize customer data in Tableau and arrange it in a dashboard. To support that, Tableau provides guidance in form of several features. In the following, we provide three examples in this regard. First, to load data, Tableau provides a process that guides users through selecting and cleaning the data - this feature would correspond to the guidance degree "prescribing" (i.e., establishing a set of actions or steps the user should take). Second, Tableau allows users to use natural language to interact with data (directing: offer a ranking or preselection of alternatives, which the user can inspect and select). Third, Tableau provides explanations in form of tooltips when users hover over data within the visualizations (orienting: keeping the user oriented without recommending actions). Following Ceneda et al. (2017), we define guidance as a concept that aims to actively resolve a skill gap encountered by users during an BI&A task.

3 Research Method

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

3.1 Derivation of the Search String

We developed our search string in several steps. We selected WebOfScience (WoS), IEEE Library, ACM DL, and Scopus for our SLR as these databases are well established and used by scholars as reliable sources for literature reviews (e.g. Bandara et al., 2015; Llave, 2017). An initial exploratory search was conducted on WoS, IEEE library, ACM DL, Scopus and Google Scholar using the search term *"Guidance AND (Business Intelligence OR Analytics)"*. After reviewing the results in several iterations, the final search string consisted of four parts: **First**, we extracted the foundational terms *"Business Intelligence"* and *"Analytics"* from our research question. We also used a related term to *"Business Intelligence"*, namely *"Decision Support Systems"*. **Second**, following Morana et al. (2017), we added two further relevant terms, namely *"explanations* and *"decision aids"*. In addition, we included the terms *"assistance"*, *"recommendation"*, and *"visual cue"* to the search string, as suggested by (Ceneda et al., 2017). We also inserted the term *"self-service"*, following Alpar and M. Schulz (2016). **Third**, according to Gregor and Benbasat (1999), the distinguishing feature of intelligent systems is "that they commonly contain a knowledge component – a computerized version of human tacit and explicit knowledge. Such systems are based on the basic elements of artificial intelligence" (Gregor and Benbasat, 1999, p. 498). For the intelligence terms of the search string, we relied on two terms provided by Serban et al. (2013), systems using*rule-based* approaches, and *machine learning*. Rule-based approaches "apply rules defined by human experts to suggest useful techniques" (Serban et al., 2013, 31:8) while systems using machine learning "automatically learn such rules from prior data analysis runs" (Serban et al., 2013, 31:8). Besides, we added the term *"knowledge-based"* to the search string (Sulaiman et al., 2015). Further, we complemented *"machine learning"*, with the term *"artificial intelligence"*, following Kühl et al. (2019). **Fourth**, for the BI&A phases, we relied on the stages defined by Chapman et al. (2000). Therefore, we added *"business understanding"*, *"data understanding"*, *"data preparation"*, *"data modeling"*, *"model evaluation"*, and *"model deployment"* including their relevant sub-phases to the search string. We also relied on the terms *"data wrangling"* and *"data integration"* for our search string as they represent an important activity for sharpening the data understanding and preparing the data for further analysis (Llave, 2018). As suggested by Bani-Hani et al. (2019), we further added the terms *"data visualization"*, *"data consumption"*, and *"data interpretation"* to the search string in order to cover activities regarding the visualization of data, consumption of created solutions, and interpretation of information provided. **Finally**, we used boolean-operators and wildcards to create the final search string:

(*("Business Intelligence" OR "Analytics" OR "Decision Support System*") AND ("Guid*" OR "Explanation*" OR "Decision Aid*" OR "Visual Cue*" OR "Self-Service" OR "Assistan*" OR "Recommendation*") AND ("Expert System" OR "Rule-based" OR "Knowledge-based" OR "Machine Learning" OR "Artificial Intelligence") AND ("Business Understanding" OR "Data Understanding" OR "Data Discovery" OR "Data Exploration" OR "Data Preparation" OR "Data Cleaning" OR "Data Integration" OR "Data Wrangling" OR "Data Selection" OR "Data Modeling" OR "Data Mining" OR "Model Evaluation" OR "Model Deployment" OR "Data Visualization" OR "Data Interpretation" OR "Data Consumption")*)

3.2 Literature Search Process

We used the defined search string and found 275 studies for WoS, 490 for IEEE, 426 for ACM DL, and 1016 for Scopus (2207 in total). We have not limited our SLR by year or publication outlet to offer a holistic overview. In the selection process, we manually excluded 1983 articles by carefully scanning the title-, abstract-, and keyword-section and by applying the following selection criteria: We included articles that covered the design, implementation, or evaluation of a BI&A system that offer guidance in at least one of the BI&A phases. We further excluded articles that addressed a system that is fully automated and thus has no options for the user to intervene, as suggested by Ceneda et al. (2017). On this basis, 224 articles were left. Following the same criteria for a full-text review, 61 relevant articles remained. Lastly, we employed a forward and backward search and included another 21 studies (in total, 82).

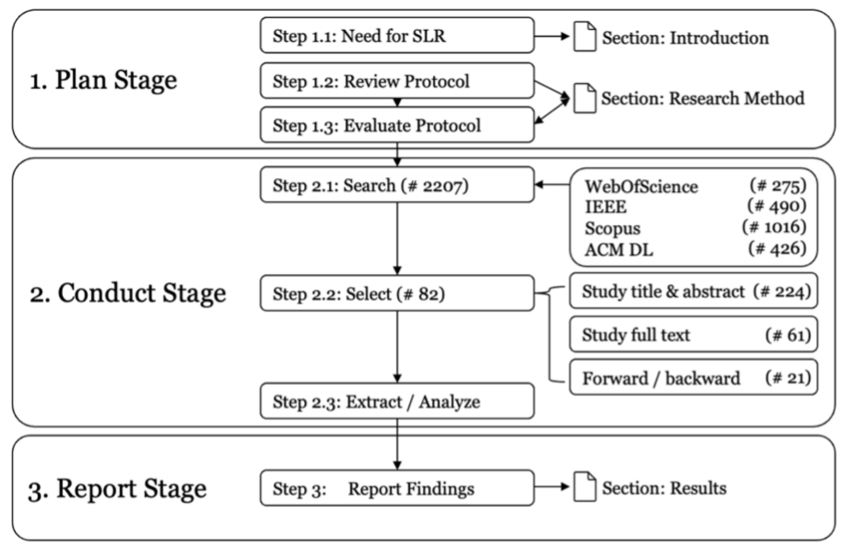

Figure 1. Literature Search Process.

3.3 Classification Process

We relied on the elements of system usage by Burton-Jones and Straub (2006) of (1) **system**, (2) **task**, and (3) **user** to set up a solid basis for the classification of guidance in BI&A systems. Building onto this, for the category of **system**, we inductively derived *interactivity form*. Specifically, we coded the *interactivity form* used in the articles along with *mouse and keyboard* and *natural language*.

For the category of **task**, we deductively derived from existing literature (i.e. Bani-Hani et al., 2019; Cao, 2017) in which *BI&A phase* the given article is offering guidance, namely *data gathering*, *data preparation*, *analytics*, *visualization* and *interpretation*. Hereby, the BI&A phases represent an abstraction of the underlying **tasks** (i.e., identify, prepare, visualize data) executed by the user.

For the category of **user**, we relied on *user roles* (deductively derived from existing literature). Specifically, we relied on the *user roles* of *business user*, *business analyst*, *data analyst*, and *data scientist* from Michalczyk et al. (2021b), which cover all user roles mentioned in the reviewed articles of this SLR.

Finally, we identified two categories for the concept of guidance, namely *guidance degree* and *guidance generation* (both deductively derived from existing literature). In particular, we followed the recommendations from Ceneda et al. (2017) and relied on three different *guidance degrees*, namely *orienting* (keeping the user oriented, without recommending actions), *directing* (offering a ranking or preselection of alternatives), and *prescribing* (including actions or step-by-step instructions the user should take). Moreover, the guidance provided was either generated top-down by human experts or were created bottom-up using an ML approach. Therefore, we used the concepts *top-down* and *bottom-up* from Collins et al. (2018) to classify the *guidance generation*.

In general, the studies were independently assigned by two researchers, and any discrepancy between the two coders' findings was discussed and resolved by consulting a third researcher (O'Connor and Joffe, 2020). For example, differences occurred in the *user roles* category, as the targeted roles were sometimes not clearly defined in the respective article. We resolved this issue by examining which user terms were specifically mentioned and how the tasks and skills of the given users were described in the article. Some differences also occurred in the *guidance degree* category, as its description was sometimes abstract. We solved these issues by considering the visual depictions contained in the articles in form of architectures or graphical user interfaces. Finally, it should be noted that the classifications are not mutually exclusive, since a given article might refer to multiple categories.

3.4 Overview of Meta-Information

The articles' publication dates (see Figure 2, left) show that the first systems were published around 2006 and peaked in 2021 (24 systems). Most systems were published as research papers in conference proceedings, followed by journal publications (see Figure 2, right). We additionally coded if the systems were evaluated following a qualitative (62% of articles) or quantitative (11%) evaluation approach. We found articles relying on interviews (e.g. Chatzimparmpas et al., 2022; Y. Wang et al., 2020) and case studies (e.g. Bian et al., 2021; Islam et al., 2021). Only a scarce amount of quantitative evaluations were conducted (Hu et al., 2018; Kretzer et al., 2015, e.g.).

4 Results

We created a concept matrix in the style suggested by Webster and Watson (2002). The complete concept matrix can be found in the Appendix. The following sections provide a results overview and emphasize the most prominent articles within each category.

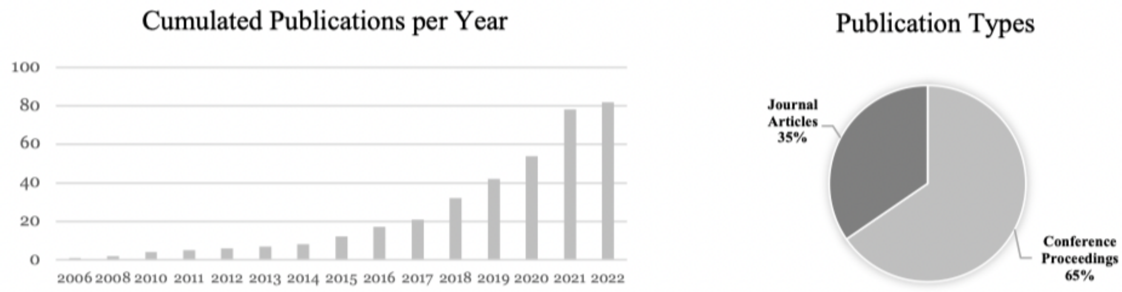

Figure 2. Cumulated Publications per Year (left) and Publication Types (right)

4.1 BI&A phases

In general, we categorized BI&A systems into five different phases according to Bani-Hani et al. (2019) and Cao (2017): (1) data gathering (26%), (2) data preparation (16%), (3) analytics (50%), (4) visualization (12%), and (5) interpretation (15%). These phases serve as a level of abstraction for the specific tasks BI&A systems support.

BI&A phases	Studies	Description	Tasks	Example Articles
Data Gathering	21	Gain general insights for further steps in the data analysis process	Access data Identify data Verify data quality	Spahn et al. (2008), Simud et al. (2020) Cui et al. (2019), Ç. Demiralp et al. (2017) Afzal et al. (2021), Kandel et al. (2012a)
Data Preparation	13	All activities to construct the final dataset from raw data	Prepare data Clean data Adjust data	Sulaiman and Gómez (2018), Mandamadiotis et al. (2021) Bourcevet et al. (2019), A. Santos et al. (2019) Afzal et al. (2021), Krause et al. (2016)
Analytics	42	Enable an understanding and discovery of actionable insight into data	Descriptive analytics Apply analytics techniques Evaluate analytics results	Cavallo and C. Demiralp (2019), Ç. Demiralp et al. (2017) Ferrettini et al. (2021), A. Santos et al. (2019) Kahng et al. (2018), A. X. Zhang et al. (2020)
Visualization	10	Visualizing and presenting the data to increase knowledge about the data	Create visualizations Create dashboards Present insights	A. Wu et al. (2022), Joshi et al. (2020) Kretzer et al. (2015) Shi et al. (2021), Islam et al. (2021)
Interpretation	9	Covers usage of existing reports, analysis, and systems for decision-making	Decision-Making	P. Xu et al. (2017), W. Wu et al. (2018)

Table 1. BI&A phases including tasks and example articles

4.1.1 Data Gathering

The main goal of this phase is to collect all relevant data and gain general insights about the data that might be helpful for the further steps in the data analysis process (Hu et al., 2018). Specific tasks cover data collection, data exploration, and verification of data quality (Cui et al., 2019). Good data quality is a prerequisite for building efficient ML models or dashboards (Afzal et al., 2021). To study and improve data quality, Kandel et al. (2012b) developed a system called Profiler that shows visualizations to help identify data quality anomalies. In turn, Afzal et al. (2021) created a system that provides visual and textual explanations to make it easier for users to explore and understand data issues flagged by the system during data profiling. Moreover, we identified various articles that helped users quickly gain insights from large datasets using data visualizations (e.g. Cui et al., 2019; Ç. Demiralp et al., 2017; Hu et al., 2018; Siddiqui et al., 2018). Other systems were designed to support more specific tasks, such as exploring causal data relationships (Xie et al., 2021), confirmatory analysis of datasets based on hypotheses (Koonchanok et al., 2021), or studying relationships between variables in multi-dimensional datasets (Knittel et al., 2021).

4.1.2 Data Preparation

Data preparation involves a series of processes that begin after ingesting raw, structured, and unstructured data (Stodder, 2016). These processes focus on improving the quality and completeness of the data, standardizing how the data is defined, consolidating the data, and finally transforming the data to make it useful for reporting and analysis (Stodder, 2016). Sulaiman and Gómez (2018) relied on analysis paths of expert users, defined as a sequence of data preparation steps, and recommended them to less-skilled users using collaborative filtering. Spahn et al. (2008) developed an ontology browser with visual query creation functionalities based on technical metadata. Further, Mandamadiotis et al. (2021) proposed to create SQL queries using natural language input and provide recommendations to select from. Moreover, the steps required to transform and prepare the data depend on whether the data will later be used to train ML models or create dashboards, so data preparation is often an integral part of the data visualization tool or done with Python (Li et al., 2021). Hereby, SeekAView represents a system that assists users in preparing data for the training of ML models by creating subspaces from high-dimensional datasets (Krause et al., 2016).

4.1.3 Analytics

This phase aims to enable an in-depth understanding and discovery of actionable insight into data. The analytics phase consists of implicit and explicit analytics. Explicit analytics focuses on descriptive analytics – typical analytics approaches consist of reporting, statistical analysis, alerting, and forecasting. While the era of implicit analytics focuses on deep analytics – typical analytics approaches are predictive modeling, optimization, prescriptive analytics, and actionable knowledge delivery (Cao, 2017). Key issues refer to (1) the selection and application of analytics techniques as well as (2) the evaluation of the outcomes (Das and Endert, 2020). Regarding the first key issue, we identified several articles that provided support for clustering data, helping users to group similar items into distinct partitions using visualizations (e.g. Cavallo and C. Demiralp, 2019; Garg et al., 2010; Kwon et al., 2018; Nauck et al., 2006). INFUSE (Krause et al., 2014) is a visual analytics tool that helps users understand the influence of the features used in their models and ranks them according to their predictive power. To support users in the feature engineering process, Chatzimparmpas et al. (2022) proposed a visual analytics system that offers multiple automatic feature selection methods. Moreover, several systems focus on automating the process of ML model building as much as possible (A. Santos et al., 2019). For instance, Ferrettini et al. (2021) recommended possible ML models based on the input data. Model selection is non-trivial: in real-world use cases, many of the best-performing models may appear to match the user goals but often exhibit nuances and tradeoffs (Sun et al., 2020). Bogl et al. (2013) provided business users with visual guidance for model selection in time series data. In addition, DFSeer (Sun et al., 2020) is an interactive visualization system for performing reliable model selection for demand forecasting based on products with similar historical demand.

The second key issue includes evaluating the trained model, inspecting parameter settings, and examining the model fitness (Chapman et al., 2000). As complex ML systems become more prevalent, it is increasingly difficult for users to understand the models or interpret the results obtained from the models (Kahng et al., 2016). Without good models and the right tools to interpret them, data scientists risk making decisions based on hidden biases and false generalizations (Hohman et al., 2019). We identified systems covering the evaluation of deep neural networks (Kahng et al., 2018; Nauta et al., 2020; Y. Wang et al., 2018; A. X. Zhang et al., 2020), random forests (Zhao et al., 2019), time series forecasting (Xie et al., 2021), and uncertain graphs (Sharara et al., 2011). In addition, several articles addressed the evaluation and comparison of classification results (N.-C. Chen et al., 2018; Gleicher et al., 2020; Krause et al., 2017). Additionally, Cheng et al. (2021) and Gomez et al. (2021) provided users with counterfactual explanations for ML models. Finally, Cabrera et al. (2019) developed a system that allows users to discover intersectional bias of ML models.

4.1.4 Visualization

The visualization phase involves visualizing the data and presenting the insights (Cao, 2017; Chapman et al., 2000). This includes presenting insights and creating dashboards using visualizations (Bani-Hani et al., 2019). MultiVision (A. Wu et al., 2022) supports users in creating analytical dashboards by providing recommendations based on deep learning. A similar system is Table2Charts (Zhou et al., 2021) which recommends possible visualizations based on patterns from data. In addition, Kretzer et al. (2015) developed a system that recommends related reports to improve the reuse of existing reports for new users. In turn, Calliope (Shi et al., 2021) assists users in creating visual data stories by suggesting visual narratives that can be inserted into the data story.

4.1.5 Interpretation

Finally, the interpretation phase covers the usage of existing reports, analysis, and systems to address decision-making tasks (Bani-Hani et al., 2019). For instance, Capelleveen et al. (2021) proposed a circular supply chain management system using interactive visualizations. In turn, P. Xu et al. (2017) provided visual diagnostics to assess assembly lines' performance in smart factories. Finally, W. Wu et al. (2018) focused on a visual analytics approach for the monitoring of equipment conditions in smart factories.

4.2 Guidance Degree

Ceneda et al. (2017) distinguished between three guidance degrees: (1) orienting (lowest degree, 63% of articles), (2) directing (medium degree, 28%), and (3) prescribing (highest degree, 10%). These guidance degrees offer a formal description for guidance in visual analytics and are based on van Wijk's model of visualization (J. v. Wijk, 2006).

Guidance Degree	Description	Studies	Example Articles
Orienting	System aims only at keeping the user oriented, without recommending actions	54	Bogl et al. (2013), Hohman et al. (2020), K. Xu et al. (2021)
Directing	System leads users along tasks by offering a ranking or preselection of alternatives	23	Kretzer et al. (2015), Shi et al. (2021), Sulaiman and Gómez (2018)
Prescribing	System establishes set of mandatory actions or step-by-step instructions the user should take	8	Hu et al. (2018), A. Santos et al. (2019), Zschech et al. (2020)

Table 2. Guidance Degree

Support for **orienting** is closely related to the goal of building and preserving a user's mental map - like a real map, it serves fundamental orienting tasks like path-finding, self-location, or exploration (Ceneda et al., 2017). For instance, the system of Gomez et al. (2021) provides predominantly visualizations to show metrics like bias, overfitting, and incorrect correlations. In addition, it relies on high contrast colors to highlight positive and negative values (Gomez et al., 2021). Cabrera et al. (2019) further included tooltips to provide more explanations. **Directing** approaches offer a ranking or preselection of alternatives, which the user can inspect and select (Ceneda et al., 2017). MultiVision (A. Wu et al., 2022) and Calliope (Shi et al., 2021) show a list of charts that the user can select and then use to visualize the data. Sulaiman and Gómez (2018) provided data preparation recommendations that the user can follow and apply to the data. While techniques that provide directions allow users to follow or ignore them, **prescribing** guidance approaches purposefully limit user influence to traversing a fixed path of analysis. The system Clustrophile 2 (Cavallo and C. Demiralp, 2019) guides users through the process of choosing clustering parameters and assessing the quality of different clustering results. Zschech et al. (2020) guided users through the process of selecting an ML model based on the problem description in natural language by providing step-by-step instructions.

4.3 Guidance Generation

As introduced before, we classified the guidance generation that the system relied on along two different categories: (1) Top-down approaches generate guidance based on derived requirements and existing rules (Sperrle et al., 2020), and (2) Bottom-up approaches generate guidance based on data (e.g., via ML techniques) (Collins et al., 2018).

Guidance Generation	Description	Studies	Example Articles
Top-down	Generate guidance top-down from requirements or existing rules	22	Afzal et al. (2021), Cui et al. (2019), Krause et al. (2014)
Bottom-up	Generate guidance bottom-up from data or user input	62	Supervised (Feng et al., 2021; Michalczyk et al., 2021a), unsupervised (Gomez et al., 2021; Krause et al., 2017), reinforcement learning (Dabek and Caban, 2017; Shi et al., 2021)

Table 3. Guidance Generation

Most of the articles analyzed relied on **bottom-up** approaches (76% of studies), while only 27% on **top-down** approaches. For instance, Afzal et al. (2021) used predefined visualizations and metrics to assess the data quality of datasets, whereas Bian et al. (2021) trained a deep learning model to enable the exploration of high dimensional datasets. We identified several systems relying on supervised (e.g. Feng et al., 2021; Michalczyk et al., 2021a; X. Wang et al., 2022), unsupervised (e.g. Gomez et al., 2021; Krause et al., 2017; Nauck et al., 2006), and reinforcement learning algorithms (e.g. Dabek and Caban, 2017; Shi et al., 2021; Tsai et al., 2015). Michalczyk et al. (2021a) guide buyers in the utilization of several pretrained classifier to reduce the supply base. Nauck et al. (2006) relied on Bayesian networks to support users in clustering unlabeled customer data. Shi et al. (2021) implemented a Monte Carlo tree search to generate and recommend visual data stories based on data.

4.4 Interactivity Form

The majority of BI&A systems interact with the target users by mouse and keyboard (90% of studies). Still, natural language interaction (10%) can be an additional interaction form.

Interactivity Form	Description	Studies	Example Articles
Mouse and keyboard	Input with mouse and keyboard	74	Cavallo and C. Demiralp (2019), Hu et al. (2018), Silva et al. (2021)
Natural language	Input using natural language	8	ML model selection (Zschech et al., 2020), querying data (Simud et al., 2020), and (3) data visualization (Lee et al., 2021; Mandamadiotis et al., 2021; Setlur et al., 2016)

Table 4. Interactivity Form

Natural language interfaces are gaining popularity due to their potential to democratize access to data and insights by making the interaction with data more natural and accessible for a wide range of users (Lee et al., 2021). In our SLR, we identified several systems that already offer natural language interaction for (1) ML model selection (Zschech et al., 2020), (2) querying data (Simud et al., 2020), and (3) data visualization (Lee et al., 2021; Mandamadiotis et al., 2021; Setlur et al., 2016). The system developed by Zschech et al. (2020) is of particular interest as it offers natural language input where users can describe their ML use case, including desired output, and the system recommends appropriate ML models.

4.5 User Roles

BI&A systems can address diverse user roles. Michalczyk et al. (2021b) distinguished between four user roles that were relevant for our classification: (1) business users (17% of articles), (2) business analysts (33%), (3) data analysts (34%), and (4) data scientists (38%). According to Lismont et al. (2019), an essential factor in the successful implementation of BI&A solutions is whether a system is capable of supporting the skills of the targeted user roles.

User Roles	Description	Studies	Example Articles
Business Users	Business domain knowledge but lack analytical skills	14	Feng et al. (2021), Islam et al. (2021), W. Wu et al. (2018)
Business Analysts	Business domain knowledge with some analytical skills	27	Da Col et al. (2021), Joshi et al. (2020), Kretzer et al. (2015)
Data Analysts	Less domain knowledge but more analytical skills	28	Cui et al. (2019), Ç. Demiralp et al. (2017), Krause et al. (2016)
Data Scientist	No specific domain knowledge but strong analytical skills	31	Garg et al. (2010), Hohman et al. (2019), Tu et al. (2021)

Table 5. User Roles

According to Michalczyk et al. (2021b), **business users** are composed of business domain knowledge without the requirement of having analytical or technical skills. TiMoVA (Bogl et al., 2013) targets business users in time series analysis model selection. Model selection in time series analysis is a challenging task for business users as it demands a close combination of domain knowledge and human judgment (Bogl et al., 2013). We further identified articles that developed systems that support business users in decision-making tasks (e.g. Feng et al., 2021; Islam et al., 2021; A. Santos et al., 2019; Tsai et al., 2015; W. Wu et al., 2018). **Business analysts** can be characterized as analytical business users, thus having strong domain knowledge. They bridge the gap between business users and data analysts by translating business problems into data projects (Saltz and Grady, 2017). For instance, Dashbot (Da Col et al., 2021) and the system of Kretzer et al. (2015) supports business analysts in the creation of dashboards. **Data analysts** have less domain understanding than business analysts but exhibit more analytical knowledge. In order to generate insights based on data, this role requires statistical modeling techniques, including business domain knowledge. For instance, DataSite (Cui et al., 2019) and Foresight (Ç. Demiralp et al., 2017) support data analysts in data exploration. Finally, **data scientists** have no specific business domain knowledge and focus on applying existing ML algorithms to data. In this SLR, we identified several systems addressing the training or evaluation of ML models that targeted data scientists (e.g. Garg et al., 2010; Hohman et al., 2019; Tu et al., 2021).

5 Suggestion of Future Research Avenues

In this section, we offer an overview of research avenues for investigating guidance in BI&A systems. Our concept matrix illustrates what has been addressed in the articles reviewed. After quantifying the state-of-the-art research, we derived what has not been addressed. For instance, our concept matrix illustrates that only 17% of current BIA systems provide guidance for business user, which is the lowest value for a targeted user role in our review – therefore we formulated the future research avenue to explore guidance for business users when designing BIA systems. Following, we emphasize avenues for future work and discuss potential solutions to address them.

5.1 Designing BI&A systems for business users with limited analytical and technical skills

Data has become increasingly strategic for most organizations, so it is no longer feasible for business users to be experts only in their specific domain. In the long run, they need to become more analytically and technically savvy (Gartner, 2016). This shift is often summarized under the term "data democratization", which refers to the act of opening organizational data to as many employees as possible (Awasthi and George, 2020). Hence, data should be made available not only to analysts or data scientists but also to non-technical or non-specialist employees like business users (Cornelissen, 2018). Currently, several self-service tools exist to support this purpose at various BI&A phases (Cao, 2017). However there is still a need to empower business users in these phases (Lismont et al., 2019), particularly due to their limited skills (Lennerholt et al., 2021). Common obstacles remain in accessing and using data, using current self-service tools, or interpreting the resulting outcomes (Lennerholt et al., 2018). Despite these obstacles, our SLR illustrated that only 17% of current BI&A systems target the business user, which is the lowest value for a targeted user role in our review. Thus, we would like to call the scholars' attention to design BI&A systems for business users, for instance, by starting design science research projects in this regard. According to Gregor and Hevner (2013), design science research is specifically suitable to theoretically derive how to design and optimize for particular complex and indecomposable problems in research and practice. We argue therefore that design science research is one promising direction to investigate the design of BI&A systems that consider guidance.

5.2 Designing user-adaptive BI&A systems

Moreover, most systems (i.e., 63%) provide the lowest guidance degree, orienting guidance. For instance, systems concerned with evaluating ML models all rely on this guidance degree. This is not necessarily an issue, as such complex tasks often require higher degrees of autonomy from the system. However, we see a tradeoff looming between providing (1) a high guidance degree and limiting autonomy for savvy users, for instance, by offering only little room for choice versus (2) a low guidance degree with more advanced features, making the system less usable for less-skilled users. Especially commercial systems like Tableau seem to have problems bridging this tradeoff. Still, we believe that supporting multiple guidance degrees is a promising step towards empowering different user roles. It would allow the guidance degree to be adapted to the individual user's advancing skills. Specifically, as a starting point for future work, we could imagine either the user providing feedback to the system or the system analyzing the user's input to adapt to the user's skills. Hereby, the body of knowledge of adaptivity research from disciplines like human-computer interaction could be made accessible for BI&A (Todi et al., 2021).

5.3 Studying the impact of natural language input on BI&A systems

There has been a steady growth in BI&A systems and interfaces to help users perform data analysis with the aid of visualizations. However, interacting with these systems can be challenging and often requires substantial user practice to become proficient (Setlur et al., 2016). Recent advances in natural language processing and human-computer interaction, such as the Google Assistant, have demonstrated this technology's potential (Yin et al., 2019). Natural language interfaces have emerged as a new way of interacting with data and performing analytics (Setlur et al., 2016). They are growing in popularity due to their potential to make the interaction with data more natural and accessible for a broader range of users (Lee et al., 2021). Hence, this approach seems promising because users may express their questions more easily in natural language rather than translate them into system commands. However, since we only identified 10% of papers using natural language interaction in this regard, we see merit for studying the impact of natural language input on BI&A systems and thus a potential avenue for future research.

5.4 Providing quantitative evaluations of BI&A systems

The results of our SLR showed that 62% of the BI&A systems relied on a qualitative user evaluation (e.g., interviews, case studies) and only 11% on a quantitative user evaluation (e.g., experiments). The importance of qualitative evaluations seems indisputable, and they are particularly well-suited for exploration (Kaplan and Duchon, 1988). Nevertheless, we believe that it is necessary to quantitatively evaluate BI&A systems as whole or their specific functions in order to understand their impact on users, for instance, by relying on survey, performance, and/or behavioral data.

5.5 Need for an integrated guidance taxonomy

Taxonomies have a long history in the natural and social sciences and also play a central role in the information systems discipline (Kundisch et al., 2021). Object classification helps researchers and practitioners understand and analyze complex domains (Miller and Roth, 1994), and further contributes to theory building by structuring and organizing the knowledge of a field, enabling the study of relationships between concepts and the generation of hypotheses about those relationships (Nickerson et al., 2013). However, given the growth of guidance literature in recent years from the fields of visual analytics (e.g. Ceneda et al., 2017), self-service BI&A (e.g. Michalczyk et al., 2020), and information visualization (e.g. Gleicher et al., 2020), we see a need to develop an integrated taxonomy of guidance characteristics. On this basis, scholars and practitioners alike could draw on the guidance characteristics of the taxonomy to design and evaluate their underlying information systems (Ceneda et al., 2020).

6 Conclusion

As current research lacks an overview of guidance in BI&A systems, we structured and analyzed existing literature with three main contributions: First, we provided a state-of-the-art overview by conducting an SLR and identifying 82 relevant articles. Second, we offered recommendations by formulating avenues for future research to be well-directed. We are aware that our article also has limitations. Any bias in the selection of the search string might result in a bias of the reviewed articles. To reduce this probability, our SLR and search process were based on well-established methodological guidelines (e.g. Kitchenham and Charters, 2007; Webster and Watson, 2002). All choices during the plan, conduct, and report stage are made explicit. We hope that our SLR can serve as a reference to support the development of BI&A systems that take into account guidance and the characteristics to be considered in this regard.

Appendix

	BI&A Phases						Guidance Degree			Guidance Generation		Interactivity Form		User Roles			
	Data Gathering	Data Preparation	Modeling	Evaluation	Deployment	Usage	Orienting	Directing	Prescribing	Top-down	Bottom-up	Mouse and Keyboard	Natural Language	Business User	Business Analyst	Data Analyst	Data Scientist
[1] Abuzaid et al. (2021)	X						X				X	X				X	
[2] Afzal et al. (2021)	X	X					X			X		X		X	X	X	X
[3] Amer-Yahia et al. (2021)	X								X		X	X			X		
[4] Bian et al. (2021)	X						X				X	X				X	
[5] Bogl et al. (2013)			X				X			X		X		X			
[6] Bourcevet et al. (2019)		X	X	X	X				X		X	X		X			
[7] Cabrera et al. (2019)				X			X				X	X					X
[8] Cavallo & Demiralp (2019)			X						X	X		X					X
[9] Chatzimparmpas et al. (2022)		X	X					X			X	X					X
[10] Chen et al. (2018)				X			X				X	X					X
[11] Cheng et al. (2021)				X			X				X	X					X
[12] Collaris & van Wijk (2020)				X			X				X	X					X
[13] Collaris & van Wijk (2022)				X			X				X	X					X
[14] Cook et al. (2015)	X							X			X	X				X	
[15] Cui et al. (2019)	X							X		X		X			X	X	
[16] Da Col et al. (2021)					X				X		X	X			X	X	
[17] Dabek & Caban (2017)	X							X			X	X			X		
[18] Das & Endert (2020)			X					X			X	X			X		
[19] Demiralp et al. (2017)	X						X			X		X			X	X	
[20] Feng et al. (2021)						X	X				X	X		X			
[21] Ferrettini et al. (2021)		X	X					X			X	X			X		
[22] Garg et al. (2010)			X				X	X			X	X					X
[23] Gibert et al. (2010)			X					X			X	X					X
[24] Gleicher et al. (2020)				X			X			X		X					X
[25] Gomez et al. (2021)				X			X				X	X					X
[26] Hohman et al. (2019)				X			X				X	X					X
[27] Hohman et al. (2020)				X			X				X	X					X
[28] Hu et al. (2018)	X	X			X				X		X	X				X	
[29] Islam et al. (2021)						X	X				X	X		X			
[30] Joshi et al.(2020)		X			X		X	X		X	X		X		X	X	
[31] Kahng et al. (2016)				X			X			X		X					X
[32] Kahng et al. (2018)				X			X				X	X					X
[33] Kahng et al. (2019)				X			X	X		X		X					X
[34] Kandel et al. (2012)	X	X					X				X	X				X	
[35] Knittel et al. (2021)	X						X				X	X				X	
[36] Koonchanok et al. (2021)	X						X			X		X				X	X
[37] Krause et al. (2014)			X				X			X		X				X	
[38] Krause et al. (2016)		X					X				X	X				X	
[39] Krause et al. (2017)				X			X				X	X					X
[40] Kretzer et al. (2015)					X			X			X	X			X		
[41] Kumar & Singh (2020)	X						X				X	X			X	X	
[42] Kwon et al. (2018)			X					X		X		X					X
[43] Langer & Meisen (2021)	X							X			X	X			X		
[44] Lee et al. (2021)	X							X			X		X		X		
[45] Lin et al. (2018)			X				X				X	X				X	
[46] Mandamadiotis et al. (2021)	X	X						X			X		X		X	X	
[47] Michalczyk et al. (2021)						X		X			X	X			X		
[48] Murugesan et al. (2019)				X			X			X		X				X	
[49] Nauck et al. (2006)			X			X			X		X	X			X		
[50] Nauta et al. (2020)				X			X				X	X					X
[51] Sacha et al. (2018)				X			X				X	X					X
[52] Santos et al. (2018)		X	X				X			X		X				X	
[53] Santos et al. (2019)			X						X		X	X		X			
[54] Setlur et al. (2016)						X		X			X		X		X	X	
[55] Sharara et al. (2011)				X			X			X		X				X	
[56] Shi et al. (2021)					X	X		X			X	X			X		
[57] Siddiqui et al. (2018)	X	X					X			X		X				X	
[58] Silva et al. (2021)						X	X				X	X				X	
[59] Simud et al. (2020)	X						X				X		X		X		
[60] Spahn et al. (2008)	X	X					X			X		X		X	X		
[61] Sulaiman & G0mez (2018)		X						X			X	X		X	X		
[62] Sun et al. (2020)			X				X			X		X			X	X	X
[63] Tu et al. (2021)			X				X				X		X				X
[64] Van Capelleveen et al. (2021)						X		X			X	X		X			
[65] Vartak et al. (2016)	X							X			X	X				X	
[66] Wang et al. (2018)				X			X				X	X					X
[67] Wang et al. (2020)				X			X				X	X					X
[68] Wang et al. (2021)				X			X			X		X					X
[69] Wang et al. (2022)				X			X				X	X					X
[70] Wongsuphasawat et al. (2016)					X			X		X		X			X	X	
[71] Wu et al. (2018)						X	X				X	X		X			
[72] Wu et al. (2022)					X			X			X	X			X	X	
[73] Xie et al. (2021)	X						X			X		X				X	
[74] Xu et al. (2015)					X		X				X	X		X			
[75] Xu et al. (2017)						X	X				X	X		X			
[76] Xu et al. (2021)				X			X				X	X					X
[77] Tsai et al. (2015)						X	X				X	X		X			
[78] Yin et al. (2019)						X	X			X	X		X		X		
[79] Zhao et al. (2019)				X			X				X	X					X
[80] Zhao et al. (2020)				X			X				X	X					X
[81] Zhou et al. (2021)					X		X				X	X			X		
[82] Zschech et al. (2020)			X						X		X		X	X	X		
Results [in %]	26	16	21	30	12	15	63	28	10	27	76	90	10	17	33	34	38

References

Abuzaid, F., P. Kraft, S. Suri, E. Gan, E. Xu, A. Shenoy, A. Ananthanarayan, J. Sheu, E. Meijer, X. Wu, J. Naughton, P. Bailis, and M. Zaharia (2021). "DIFF: a relational interface for large-scale data explanation." *The VLDB Journal* 30 (1), 45–70.

Afzal, S., A. Chaudhary, N. Gupta, H. Patel, C. Spina, and D. Wang (2021). "Data-Debugging Through Interactive Visual Explanations." In: *Trends and Applications in Knowledge Discovery and Data Mining*. Ed. by M. Gupta and G. Ramakrishnan. Cham: Springer International Publishing, pp. 133–142.

Agrawal, S., A. De Smet, P. Poplawski, and A. Reich (2020). *Beyond hiring: How companies are reskilling to address talent gaps*.

Alpar, P. and M. Schulz (2016). "Self-Service Business Intelligence." *Business & Information Systems Engineering* 58 (2), 151–155.

Amer-Yahia, S., G. Koutrika, F. Bastian, T. Belmpas, M. Braschler, U. Brunner, D. Calvanese, M. Fabricius, O. Gkini, C. Kosten, D. Lanti, A. Litke, H. Lücke-Tieke, F. A. Massucci, T. M. de Farias, A. Mosca, F. Multari, N. Papadakis, D. Papadopoulos, Y. Patil, A. Personnaz, G. Rull, A. Sima, E. Smith, D. Skoutas, S. Subramanian, G. Xiao, and K. Stockinger (2021). "INODE: Building an End-to-End Data Exploration System in Practice [Extended Vision]." *arXiv:2104.04194 [cs]*. arXiv: 2104.04194.

Awasthi, P. and J. J. George (2020). "A case for Data Democratization." *AMCIS 2020 Proceedings. 23*, 11.

Bandara, W., E. Furtmueller, E. Gorbacheva, S. Miskon, and J. Beekhuyzen (2015). "Achieving Rigor in Literature Reviews: Insights from Qualitative Data Analysis and Tool-Support." *Communications of the Association for Information Systems* 37.

Bani-Hani, I., O. Tona, and S. Carlsson (2019). "Modes of engagement in SSBA: A service dominant logic perspective." In: Association for Information Systems.

Bian, Y., C. North, E. Krokos, and S. Joseph (2021). "Semantic Explanation of Interactive Dimensionality Reduction." In: *2021 IEEE Visualization Conference (VIS)*. New Orleans, LA, USA: IEEE, pp. 26–30.

Bogl, M., W. Aigner, P. Filzmoser, T. Lammarsch, S. Miksch, and A. Rind (2013). "Visual Analytics for Model Selection in Time Series Analysis." *IEEE Transactions on Visualization and Computer Graphics* 19 (12), 2237–2246.

Bourcevet, A., G. Piller, University of Applied Sciences Mainz, Mainz, Germany, M. Scholz, University of Applied Sciences Mainz, Mainz, Germany, J. Wiesemann, and CubeServ GmbH, Raunheim, Germany (2019). "Guided Machine Learning for Business Users." In: *Humanizing Technology for a Sustainable Society*. Univresity of Maribor Press, pp. 257–270.

Burton-Jones, A. and D. W. Straub (2006). "Reconceptualizing System Usage: An Approach and Empirical Test." *Information Systems Research* 17 (3), 228–246.

Cabrera, A. A., W. Epperson, F. Hohman, M. Kahng, J. Morgenstern, and D. H. Chau (2019). "FAIRVIS: Visual Analytics for Discovering Intersectional Bias in Machine Learning." In: *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Vancouver, BC, Canada: IEEE, pp. 46–56.

Cao, L. (2017). "Data Science: A Comprehensive Overview." *ACM Computing Surveys* 50 (3), 1–42.

Capelleveen, G. van, J. van Wieren, C. Amrit, D. M. Yazan, and H. Zijm (2021). "Exploring recommendations for circular supply chain management through interactive visualisation." *Decision Support Systems* 140, 113431.

Cavallo, M. and C. Demiralp (2019). "Clustrophile 2: Guided Visual Clustering Analysis." *IEEE Transactions on Visualization and Computer Graphics* 25 (1), 267–276.

Ceneda, D., N. Andrienko, G. Andrienko, T. Gschwandtner, S. Miksch, N. Piccolotto, T. Schreck, M. Streit, J. Suschnigg, and C. Tominski (2020). "Guide Me in Analysis: A Framework for Guidance Designers." *Computer Graphics Forum* 39 (6), 269–288.

Ceneda, D., T. Gschwandtner, T. May, S. Miksch, H.-J. Schulz, M. Streit, and C. Tominski (2017). "Characterizing Guidance in Visual Analytics." *IEEE Transactions on Visualization and Computer Graphics* 23 (1), 111–120.

Ceneda, D., T. Gschwandtner, and S. Miksch (2019). "A Review of Guidance Approaches in Visual Data Analysis: A Multifocal Perspective." *Computer Graphics Forum* 38 (3), 861–879.

Chae, B. (and D. L. Olson (2013). "BUSINESS ANALYTICS FOR SUPPLY CHAIN: A DYNAMIC-CAPABILITIES FRAMEWORK." *International Journal of Information Technology & Decision Making* 12 (01), 9–26.

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

Chatzimparmpas, A., R. M. Martins, K. Kucher, and A. Kerren (2022). "FeatureEnVi: Visual Analytics for Feature Engineering Using Stepwise Selection and Semi-Automatic Extraction Approaches." *IEEE Transactions on Visualization and Computer Graphics* 28 (4), 1773–1791.

Chaudhuri, S., U. Dayal, and V. Narasayya (2011). "An overview of business intelligence technology." *Communications of the ACM* 54 (8), 88–98.

Chen, Chiang, and Storey (2012). "Business Intelligence and Analytics: From Big Data to Big Impact." *MIS Quarterly* 36 (4), 1165.

Chen, N.-C., J. Suh, J. Verwey, G. Ramos, S. Drucker, and P. Simard (2018). "AnchorViz: Facilitating Classifier Error Discovery through Interactive Semantic Data Exploration." In: *23rd International Conference on Intelligent User Interfaces*. Tokyo Japan: ACM, pp. 269–280.

Cheng, F., Y. Ming, and H. Qu (2021). "DECE: Decision Explorer with Counterfactual Explanations for Machine Learning Models." *IEEE Transactions on Visualization and Computer Graphics* 27 (2), 1438–1447.

Collaris, D. and J. Van Wijk (2022). "StrategyAtlas: Strategy Analysis for Machine Learning Interpretability." *IEEE Transactions on Visualization and Computer Graphics*, 1–1.

Collaris, D. and J. J. van Wijk (2020). "ExplainExplore: Visual Exploration of Machine Learning Explanations." In: *2020 IEEE Pacific Visualization Symposium (PacificVis)*. Tianjin, China: IEEE, pp. 26–35.

Collins, C., N. Andrienko, T. Schreck, J. Yang, J. Choo, U. Engelke, A. Jena, and T. Dwyer (2018). "Guidance in the human–machine analytics process." *Visual Informatics* 2 (3), 166–180.

Cook, K., N. Cramer, D. Israel, M. Wolverton, J. Bruce, R. Burtner, and A. Endert (2015). "Mixed-initiative visual analytics using task-driven recommendations." In: *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Chicago, IL, USA: IEEE, pp. 9–16.

Cornelissen, J. (2018). "The Democratization of Data Science." *Harvard Business Review*. Section: Analytics and data science.

Cui, Z., S. K. Badam, M. A. Yalçin, and N. Elmqvist (2019). "DataSite: Proactive visual data exploration with computation of insight-based recommendations." *Information Visualization* 18 (2), 251–267.

Da Col, S., R. Ciucanu, M. Soare, N. Bouarour, and S. Amer-Yahia (2021). "DashBot: An ML-Guided Dashboard Generation System." In: *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. Virtual Event Queensland Australia: ACM, pp. 4696–4700.

Dabek, F. and J. J. Caban (2017). "A Grammar-based Approach for Modeling User Interactions and Generating Suggestions During the Data Exploration Process." *IEEE Transactions on Visualization and Computer Graphics* 23 (1), 41–50.

Das, S. and A. Endert (2020). "LEGION: Visually compare modeling techniques for regression." In: *2020 Visualization in Data Science (VDS)*. Salt Lake City, UT, USA: IEEE, pp. 12–21.

Demiralp, Ç., P. J. Haas, S. Parthasarathy, and T. Pedapati (2017). "Foresight: Recommending Visual Insights." *arXiv:1707.03877 [cs]*. arXiv: 1707.03877.

Feng, Y., J. Xu, W. Sheng, J. Chen, and Y. Hong (2021). "Developing a Decision Support System for Water Resources Dispatching." In: *Frontiers in Artificial Intelligence and Applications*. IOS Press.

Ferrettini, G., J. Aligon, C. Soulé-Dupuy, and W. Raynaut (2021). "A framework for user assistance on predictive models." *Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020). Vol. 2621.*, 9.

Garg, S., I. V. Ramakrishnan, and K. Mueller (2010). "A visual analytics approach to model learning." In: *2010 IEEE Symposium on Visual Analytics Science and Technology*. Salt Lake City, UT, USA: IEEE, pp. 67–74.

Gartner (2016). *Gartner Says Worldwide Business Intelligence and Analytics Market to Reach $16.9 Billion in 2016*.

Gibert, K., M. Sànchez-Marrè, and V. Codina (2010). "Choosing the Right Data Mining Technique: Classification of Methods and Intelligent Recommendation," 10.

Gleicher, M., A. Barve, X. Yu, and F. Heimerl (2020). "Boxer: Interactive Comparison of Classifier Results." *Computer Graphics Forum* 39 (3), 181–193.

Gomez, O., S. Holter, J. Yuan, and E. Bertini (2021). "AdViCE: Aggregated Visual Counterfactual Explanations for Machine Learning Model Validation." *arXiv:2109.05629 [cs]*. arXiv: 2109.05629.

Gregor, S. and I. Benbasat (1999). "Explanations from Intelligent Systems: Theoretical Foundations and Implications for Practice." *MIS Quarterly* 23 (4), 497.

Gregor, S. and A. R. Hevner (2013). "Positioning and Presenting Design Science Research for Maximum Impact." *MIS Quarterly* 37 (2), 337–355.

Hohman, F., A. Head, R. Caruana, R. DeLine, and S. M. Drucker (2019). "Gamut: A Design Probe to Understand How Data Scientists Understand Machine Learning Models." In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Glasgow Scotland Uk: ACM, pp. 1–13.

Hohman, F., H. Park, C. Robinson, and D. H. Polo Chau (2020). "Summit: Scaling Deep Learning Interpretability by Visualizing Activation and Attribution Summarizations." *IEEE Transactions on Visualization and Computer Graphics* 26 (1), 1096–1106.

Hu, K., D. Orghian, and C. Hidalgo (2018). "DIVE: A Mixed-Initiative System Supporting Integrated Data Exploration Workflows." In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics*. Houston TX USA: ACM, pp. 1–7.

Huguenard, B. and M. Frolick (2001). "You May Feel Better Off Than You Are: Usefulness Evaluations of Cognitive Feedback." *AMCIS 2001 Proceedings. 56.*, 8.

Imhoff, C. and C. White (2011). "Self-Service Business Intelligence: Empowering Users to Generate Insights." *TDWI Best practices report*, 40.

Islam, M. R., J. Zhang, M. H. Ashmafee, I. Razzak, J. Zhou, X. Wang, and G. Xu (2021). "ExVis: Explainable Visual Decision Support System for Risk Management." In: *2021 8th International Conference on Behavioral and Social Computing (BESC)*. Doha, Qatar: IEEE, pp. 1–5.

Joshi, S. R., B. Venkatesh, D. Thomas, Y. Jiao, and S. Roy (2020). "A Natural Language and Interactive End-to-End Querying and Reporting System." In: *Proceedings of the 7th ACM IKDD CoDS and 25th COMAD*. Hyderabad India: ACM, pp. 261–267.

Kahng, M., P. Y. Andrews, A. Kalro, and D. H. Chau (2018). "ActiVis: Visual Exploration of Industry-Scale Deep Neural Network Models." *IEEE Transactions on Visualization and Computer Graphics* 24 (1), 88–97.

Kahng, M., D. Fang, and D. H. (Chau (2016). "Visual exploration of machine learning results using data cube analysis." In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics - HILDA '16*. San Francisco, California: ACM Press, pp. 1–6.

Kahng, M., N. Thorat, D. H. P. Chau, F. B. Viegas, and M. Wattenberg (2019). "GAN Lab: Understanding Complex Deep Generative Models using Interactive Visual Experimentation." *IEEE Transactions on Visualization and Computer Graphics* 25 (1), 310–320.

Kandel, S., A. Paepcke, J. M. Hellerstein, and J. Heer (2012a). "Enterprise Data Analysis and Visualization: An Interview Study." *IEEE Transactions on Visualization and Computer Graphics* 18 (12), 2917–2926.

Kandel, S., R. Parikh, A. Paepcke, J. M. Hellerstein, and J. Heer (2012b). "Profiler: integrated statistical analysis and visualization for data quality assessment." In: *Proceedings of the International Working Conference on Advanced Visual Interfaces - AVI '12*. Capri Island, Italy: ACM Press, p. 547.

Kaplan, B. and D. Duchon (1988). "Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study." *MIS Quarterly* 12 (4), 571.

Kitchenham, B. and S. Charters (2007). “Guidelines for performing systematic literature reviews in software engineering.” *EBSE Technical Report. EBSE-2007-01*. Publisher: Citeseer.

Knittel, J., A. Lalama, S. Koch, and T. Ertl (2021). “Visual Neural Decomposition to Explain Multivariate Data Sets.” *IEEE Transactions on Visualization and Computer Graphics* 27 (2), 1374–1384.

Koonchanok, R., P. Baser, A. Sikharam, N. K. Raveendranath, and K. Reda (2021). “Data Prophecy: Exploring the Effects of Belief Elicitation in Visual Analytics.” In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, pp. 1–12.

Krause, J., A. Dasgupta, J.-D. Fekete, and E. Bertini (2016). “SeekAView: An intelligent dimensionality reduction strategy for navigating high-dimensional data spaces.” In: *2016 IEEE 6th Symposium on Large Data Analysis and Visualization (LDAV)*. Baltimore, MD, USA: IEEE, pp. 11–19.

Krause, J., A. Dasgupta, J. Swartz, Y. Aphinyanaphongs, and E. Bertini (2017). “A Workflow for Visual Diagnostics of Binary Classifiers using Instance-Level Explanations.” In: *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Phoenix, AZ: IEEE, pp. 162–172.

Krause, J., A. Perer, and E. Bertini (2014). “INFUSE: Interactive Feature Selection for Predictive Modeling of High Dimensional Data.” *IEEE Transactions on Visualization and Computer Graphics* 20 (12), 1614–1623.

Kretzer, M., M. Kretzer, M. Kleinedler, C. Theilemann, and A. Mädche (2015). “Designing a Report Recommendation Assistant: A First Design Cycle.” In: *New Horizons in Design Science: Broadening the Research Agenda*. Vol. 9073. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 87–103.

Kühl, N., M. Goutier, R. Hirt, and G. Satzger (2019). “Machine Learning in Artificial Intelligence: Towards a Common Understanding.” In.

Kumar, V. and V. Singh (2020). “An Automatic Intent Modeling Algorithm for Interactive Data Exploration.” In: *Computational Intelligence in Data Mining*. Ed. by H. S. Behera, J. Nayak, B. Naik, and D. Pelusi. Singapore: Springer Singapore, pp. 127–140.

Kundisch, D., J. Muntermann, A. M. Oberländer, D. Rau, M. Röglinger, T. Schoormann, and D. Szopinski (2021). “An Update for Taxonomy Designers: Methodological Guidance from Information Systems Research.” *Business & Information Systems Engineering*.

Kwon, B. C., B. Eysenbach, J. Verma, K. Ng, C. De Filippi, W. F. Stewart, and A. Perer (2018). “Clustervision: Visual Supervision of Unsupervised Clustering.” *IEEE Transactions on Visualization and Computer Graphics* 24 (1), 142–151.

Langer, T. and T. Meisen (2021). “System Design to Utilize Domain Expertise for Visual Exploratory Data Analysis.” *Information* 12 (4), 140.

Lee, D. J.-L., A. Quamar, E. Kandogan, and F. Özcan (2021). “Boomerang: Proactive Insight-Based Recommendations for Guiding Conversational Data Analysis.” In: *Proceedings of the 2021 International Conference on Management of Data*. Virtual Event China: ACM, pp. 2750–2754.

Lennerholt, C., J. van Laere, and E. Söderström (2018). “Implementation Challenges of Self Service Business Intelligence: A Literature Review.” *51st Hawaii International Conference on System Sciences*, 9.

Lennerholt, C., J. Van Laere, and E. Söderström (2021). “User-Related Challenges of Self-Service Business Intelligence.” *Information Systems Management* 38 (4), 309–323.

Lexico (2022). *Guidance*.

Li, X., Y. Zhang, J. Leung, C. Sun, and J. Zhao (2021). “EDAssistant: Supporting Exploratory Data Analysis in Computational Notebooks with In-Situ Code Search and Recommendation.” *arXiv:2112.07858 [cs]*. arXiv: 2112.07858.

Lim, E.-P., H. Chen, and G. Chen (2013). “Business Intelligence and Analytics: Research Directions.” *ACM Transactions on Management Information Systems* 3 (4), 1–10.

Lin, H., S. Gao, D. Gotz, F. Du, J. He, and N. Cao (2018). “RCLens: Interactive Rare Category Exploration and Identification.” *IEEE Transactions on Visualization and Computer Graphics* 24 (7), 2223–2237.

Lismont, J., T. Van Calster, M. Óskarsdóttir, S. vanden Broucke, B. Baesens, W. Lemahieu, and J. Vanthienen (2019). "Closing the Gap Between Experts and Novices Using Analytics-as-a-Service: An Experimental Study." *Business & Information Systems Engineering* 61 (6), 679–693.

Llave, M. R. (2017). "Business Intelligence and Analytics in Small and Medium-sized Enterprises: A Systematic Literature Review." *Procedia Computer Science* 121, 194–205.

— (2018). "Data lakes in business intelligence: reporting from the trenches." *Procedia Computer Science* 138, 516–524.

Mandamadiotis, A., S. Eleftherakis, A. Glenis, D. Skoutas, Y. Stavrakas, and G. Koutrika (2021). "DatAgent: the imminent age of intelligent data assistants." *Proceedings of the VLDB Endowment* 14 (12), 2815–2818.

Merriam-Webster (2022). *Guidance*.

Michalczyk, S., M. Nadj, D. Azarfar, A. Maedche, and C. Gröger (2020). "A State-of-the-Art Overview and Future Research Avenues of Self-Service Business Intelligence & Analytics." *European Conference on Information Systems*, 21.

Michalczyk, S., M. Nadj, H. Beier, and A. Maedche (2021a). "Designing a Self-service Analytics System for Supply Base Optimization." In: *Intelligent Information Systems*. Ed. by S. Nurcan and A. Korthaus. Cham: Springer International Publishing, pp. 154–161.

Michalczyk, S., M. Nadj, A. Maedche, and C. Gröger (2021b). "Demystifying Job Roles in Data Science: A Text Mining Approach." *ECIS 2021 Research Papers*.

Miller, J. G. and A. V. Roth (1994). "A Taxonomy of Manufacturing Strategies." *Management Science* 40 (3), 285–304.

Miloslavskaya, N. and A. Tolstoy (2016). "Big Data, Fast Data and Data Lake Concepts." *Procedia Computer Science* 88, 300–305.

Morana, S., S. Schacht, A. Scherp, and A. Maedche (2017). "A review of the nature and effects of guidance design features." *Decision Support Systems* 97, 31–42.

Murugesan, S., S. Malik, F. Du, E. Koh, and T. M. Lai (2019). "DeepCompare: Visual and Interactive Comparison of Deep Learning Model Performance." *IEEE Computer Graphics and Applications* 39 (5), 47–59.

Nauck, D. D., D. Ruta, M. Spott, and B. Azvine (2006). "A Tool for Intelligent Customer Analytics." In: *2006 3rd International IEEE Conference Intelligent Systems*. London, UK: IEEE, pp. 518–521.

Nauta, M., J. P. Bos, M. van Keulen, and C. Seifert (2020). "Interactive Explanations of Internal Representations of Neural Network Layers: An Exploratory Study on Outcome Prediction of Comatose Patients," 7.

Nickerson, R. C., U. Varshney, and J. Muntermann (2013). "A method for taxonomy development and its application in information systems." *European Journal of Information Systems* 22 (3), 336–359.

O'Connor, C. and H. Joffe (2020). "Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines." *International Journal of Qualitative Methods* 19, 160940691989922.

Rikhardsson, P. and O. Yigitbasioglu (2018). "Business intelligence & analytics in management accounting research: Status and future focus." *International Journal of Accounting Information Systems* 29, 37–58.

Sacha, D., M. Kraus, J. Bernard, M. Behrisch, T. Schreck, Y. Asano, and D. A. Keim (2018). "SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance." *IEEE Transactions on Visualization and Computer Graphics* 24 (1), 120–130.

Saltz, J. S. and N. W. Grady (2017). "The ambiguity of data science team roles and the need for a data science workforce framework." In: *2017 IEEE International Conference on Big Data (Big Data)*. Boston, MA: IEEE, pp. 2355–2361.

Santos, A., S. Castelo, C. Felix, J. P. Ono, B. Yu, S. R. Hong, C. T. Silva, E. Bertini, and J. Freire (2019). "Visus: An Interactive System for Automatic Machine Learning Model Building and Curation." In: *Proceedings of the Workshop on Human-In-the-Loop Data Analytics - HILDA'19*. Amsterdam, Netherlands: ACM Press, pp. 1–7.

Santos, W. d., G. P. Avelar, M. H. Ribeiro, D. Guedes, and W. Meira (2018). "Scalable and efficient data analytics and mining with lemonade." *Proceedings of the VLDB Endowment* 11 (12), 2070–2073.

Schulz, H.-J., M. Streit, T. May, and C. Tominski (2013). "Towards a characterization of guidance in visualization." In.

Serban, F., J. Vanschoren, J.-U. Kietz, and A. Bernstein (2013). "A survey of intelligent assistants for data analysis." *ACM Computing Surveys* 45 (3), 1–35.

Setlur, V., S. E. Battersby, M. Tory, R. Gossweiler, and A. X. Chang (2016). "Eviza: A Natural Language Interface for Visual Analysis." In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. Tokyo Japan: ACM, pp. 365–377.

Sharara, H., A. Sopan, G. Namata, L. Getoor, and L. Singh (2011). "G-PARE: A visual analytic tool for comparative analysis of uncertain graphs." In: *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*. Providence, RI, USA: IEEE, pp. 61–70.

Shen, M., M. Carswell, R. Santhanam, and K. Bailey (2012). "Emergency management information systems: Could decision makers be supported in choosing display formats?" *Decision Support Systems* 52 (2), 318–330.

Shi, D., X. Xu, F. Sun, Y. Shi, and N. Cao (2021). "Calliope: Automatic Visual Data Story Generation from a Spreadsheet." *IEEE Transactions on Visualization and Computer Graphics* 27 (2), 453–463.

Siddiqui, T., A. Kim, J. Lee, K. Karahalios, and A. Parameswaran (2018). "Effortless Data Exploration with zenvisage: An Expressive and Interactive Visual Analytics System." *arXiv:1604.03583 [cs]*. arXiv: 1604.03583.

Silva, P., C. Macas, E. Polisciuc, and P. Machado (2021). "Visualisation Tool to Support Fraud Detection." In: *2021 25th International Conference Information Visualisation (IV)*. Sydney, Australia: IEEE, pp. 77–87.

Silver, M. S. (1991). "Decisional Guidance for Computer-Based Decision Support." *MIS Quarterly* 15 (1), 105.

Simud, T., S. Ruengittinun, N. Surasvadi, N. Sanglerdsinlapachai, and A. Plangprasopchok (2020). "A Conversational Agent for Database Query: A Use Case for Thai People Map and Analytics Platform." In: *2020 15th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP)*. Bangkok, Thailand: IEEE, pp. 1–6.

Smith, S. L. and J. N. Mosier (1986). *Guidelines for designing user interface software*. Citeseer.

Spahn, M., J. Kleb, S. Grimm, and S. Scheidl (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 '08*. Karlsruhe, Germany: ACM Press, pp. 1–12.

Sperrle, F., A. Jeitler, J. Bernard, D. A. Keim, and M. El-Assady (2020). "Learning and Teaching in Co-Adaptive Guidance for Mixed-Initiative Visual Analytics." *EuroVis Workshop on Visual Analytics (EuroVA)*. Artwork Size: 5 pages ISBN: 9783038681168 Publisher: The Eurographics Association Version Number: 061-065, 5 pages.

Stodder, D. (2016). "Improving data preparation for business analytics." *Transforming Data With Intelligence* 1 (1), 41.

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

Sulaiman, S., T. Mahmoud, J. Marx, and J. Kurzhöfer (2015). "Automatic Knowledge Transfer-based Architecture towards Self-Service Business Intelligence." *DBKDA 2015 . The Seventh International Conference on Advances in Databases, Knowledge, and Data Applications*, 8.

Sun, D., Z. Feng, Y. Chen, Y. Wang, J. Zeng, M. Yuan, T.-C. Pong, and H. Qu (2020). "DFSeer: A Visual Analytics Approach to Facilitate Model Selection for Demand Forecasting." In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Honolulu HI USA: ACM, pp. 1–13.

Todi, K., G. Bailly, L. A. Leiva, and A. Oulasvirta (2021). "Adapting User Interfaces with Model-based Reinforcement Learning." *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. arXiv: 2103.06807, 1–13.

Trieu, V.-H., A. Burton-Jones, The University of Queensland, P. Green, Queensland University of Technology, S. Cockcroft, and University of Southern Queensland, Springfield Campus (2022). "Applying and Extending the Theory of Effective Use in a Business Intelligence Context." *MIS Quarterly* 46 (1), 645–678.

Tsai, Y.-C., L.-C. Hsieh, W.-F. Cheng, Y.-H. Kuo, W. Hsu, and W.-C. Chen (2015). "Trending pool: Visual analytics for trending event compositions for time-series categorical log data." In: *2015 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 221–222.

Tu, Y., J. Xu, and H.-W. Shen (2021). "KeywordMap: Attention-based Visual Exploration for Keyword Analysis." In: *2021 IEEE 14th Pacific Visualization Symposium (PacificVis)*. Tianjin, China: IEEE, pp. 206–215.

Vailshery, L. S. (2022). *BI & analytics software market value worldwide 2019-2025*.

Vartak, M., S. Rahman, S. Madden, and A. Parameswaran (2016). "SEEDB: Efficient Data-Driven Visualization Recommendations to Support Visual Analytics," 41.

Wang, J., W. Zhang, and H. Yang (2020). "SCANViz: Interpreting the Symbol-Concept Association Captured by Deep Neural Networks through Visual Analytics." In: *2020 IEEE Pacific Visualization Symposium (PacificVis)*. Tianjin, China: IEEE, pp. 51–60.

Wang, Q., Z. Xu, Z. Chen, Y. Wang, S. Liu, and H. Qu (2021). "Visual Analysis of Discrimination in Machine Learning." *IEEE Transactions on Visualization and Computer Graphics* 27 (2), 1470–1480.

Wang, X., J. He, Z. Jin, M. Yang, Y. Wang, and H. Qu (2022). "M2Lens: Visualizing and Explaining Multimodal Models for Sentiment Analysis." *IEEE Transactions on Visualization and Computer Graphics* 28 (1), 802–812.

Wang, Y., Z. Sun, H. Zhang, W. Cui, K. Xu, X. Ma, and D. Zhang (2020). "DataShot: Automatic Generation of Fact Sheets from Tabular Data." *IEEE Transactions on Visualization and Computer Graphics* 26 (1), 895–905.

Wang, Y., H. Zhang, H. Huang, X. Chen, Q. Yin, Z. Hou, D. Zhang, Q. Luo, and H. Qu (2018). "InfoNice: Easy Creation of Information Graphics." In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Montreal QC Canada: ACM, pp. 1–12.

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

Wijk, J. van (2006). "Views on Visualization." *IEEE Transactions on Visualization and Computer Graphics* 12 (4), 421–432.

Wongsuphasawat, K., D. Moritz, A. Anand, J. Mackinlay, B. Howe, and J. Heer (2016). "Voyager: Exploratory Analysis via Faceted Browsing of Visualization Recommendations." *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS* 22 (1), 10.

Wu, A., Y. Wang, M. Zhou, X. He, H. Zhang, H. Qu, and D. Zhang (2022). "MultiVision: Designing Analytical Dashboards with Deep Learning Based Recommendation." *IEEE Transactions on Visualization and Computer Graphics* 28 (1), 162–172.

Wu, W., Y. Zheng, K. Chen, X. Wang, and N. Cao (2018). "A Visual Analytics Approach for Equipment Condition Monitoring in Smart Factories of Process Industry." In: *2018 IEEE Pacific Visualization Symposium (PacificVis)*. Kobe: IEEE, pp. 140–149.

Xie, X., F. Du, and Y. Wu (2021). "A Visual Analytics Approach for Exploratory Causal Analysis: Exploration, Validation, and Applications." *IEEE Transactions on Visualization and Computer Graphics* 27 (2), 1448–1458.

Xu, C., T. Neuroth, T. Fujiwara, R. Liang, and K.-L. Ma (2015). "A Predictive Visual Analytics System for Studying Neurodegenerative Disease based on DTI Fiber Tracts." *arXiv:2010.07047 [cs, eess]*. arXiv: 2010.07047.

Xu, K., J. Yuan, Y. Wang, C. Silva, and E. Bertini (2021). "mTSeer: Interactive Visual Exploration of Models on Multivariate Time-series Forecast." In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Yokohama Japan: ACM, pp. 1–15.

Xu, P., H. Mei, L. Ren, and W. Chen (2017). "ViDX: Visual Diagnostics of Assembly Line Performance in Smart Factories." *IEEE Transactions on Visualization and Computer Graphics* 23 (1), 291–300.

Ye, L. R. and P. E. Johnson (1995). "The Impact of Explanation Facilities on User Acceptance of Expert Systems Advice." *MIS Quarterly* 19 (2), 157.

Yin, Z., C. Zhang, D. W. Goldberg, and S. Prasad (2019). "An NLP-based Question Answering Framework for Spatio-Temporal Analysis and Visualization." In: *Proceedings of the 2019 2nd International Conference on Geoinformatics and Data Analysis*. Prague Czech Republic: ACM, pp. 61–65.

Zhang, A. X., M. Muller, and D. Wang (2020). "How do Data Science Workers Collaborate? Roles, Workflows, and Tools." *Proceedings of the ACM on Human-Computer Interaction* 4 (CSCW1), 1–23.

Zhang, Y. and A. Lugmayr (2019). "Designing a User-Centered Interactive Data-Storytelling Framework." In: *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*. Fremantle WA Australia: ACM, pp. 428–432.

Zhao, X., Y. Wu, D. L. Lee, and W. Cui (2019). "iForest: Interpreting Random Forests via Visual Analytics." *IEEE Transactions on Visualization and Computer Graphics* 25 (1), 407–416.

Zhou, M., Q. Li, X. He, Y. Li, Y. Liu, W. Ji, S. Han, Y. Chen, D. Jiang, and D. Zhang (2021). "Table2Charts: Recommending Charts by Learning Shared Table Representations." In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*. Virtual Event Singapore: ACM, pp. 2389–2399.

Zschech, P., R. Horn, D. Höschele, C. Janiesch, and K. Heinrich (2020). "Intelligent User Assistance for Automated Data Mining Method Selection." *Business & Information Systems Engineering* 62 (3), 227–247.