Characterizing Guidance in Visual Analytics

Davide Ceneda, Theresia Gschwandtner, Thorsten May,
Silvia Miksch, Hans-Jörg Schulz, Marc Streit, and Christian Tominski

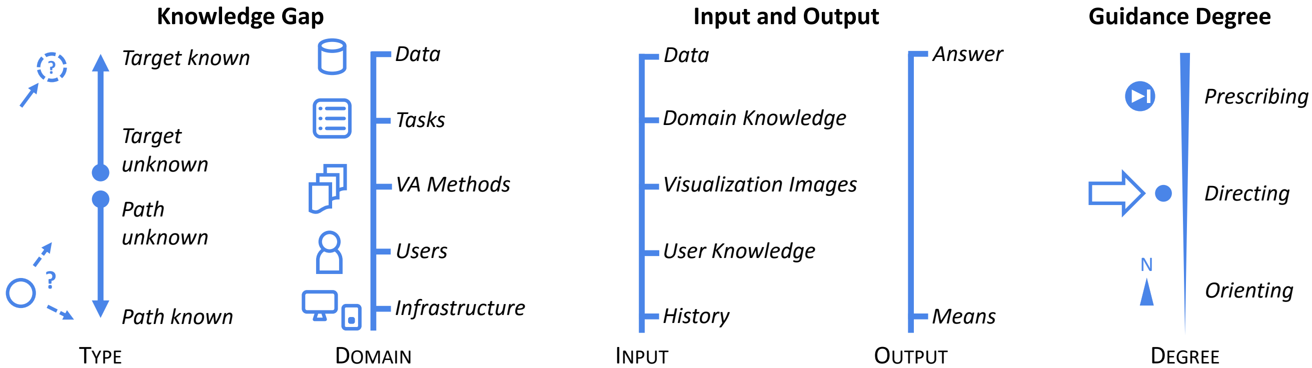


Fig. 1: Guidance can be characterized in terms of the main aspects: knowledge gap, input and output, as well as guidance degree.

Abstract—Visual analytics (VA) is typically applied in scenarios where complex data has to be analyzed. Unfortunately, there is a natural correlation between the complexity of the data and the complexity of the tools to study them. An adverse effect of complicated tools is that analytical goals are more difficult to reach. Therefore, it makes sense to consider methods that guide or assist users in the visual analysis process. Several such methods already exist in the literature, yet we are lacking a general model that facilitates in-depth reasoning about guidance. We establish such a model by extending van Wijk's model of visualization with the fundamental components of guidance. Guidance is defined as a process that gradually narrows the gap that hinders effective continuation of the data analysis. We describe diverse inputs based on which guidance can be generated and discuss different degrees of guidance and means to incorporate guidance into VA tools. We use existing guidance approaches from the literature to illustrate the various aspects of our model. As a conclusion, we identify research challenges and suggest directions for future studies. With our work we take a necessary step to pave the way to a systematic development of guidance techniques that effectively support users in the context of VA.

Index Terms—Visual analytics, guidance model, assistance, user support

1 Introduction

Thomas and Cook [46] define visual analytics (VA) as a technology that supports discovery by combining automated analysis with interactive visual means. The key idea is to establish a synergy of computational power and human reasoning. In recent years, a large number of VA approaches have been developed for diverse data, analytical problems, and user requirements. They are particularly useful in situations where complex problems have to be solved. Consequently, these methods are often not as simple to use as one would wish they were. Analytical computations usually require the user to set parameters, while suitable values are not clear upfront. Visual representations of complex phenomena tend to be more demanding to interpret than plain information graphics. And also in terms of interaction there are many more things to control, in order to make proper progress in the data analysis process.

- *Davide Ceneda, Theresia Gschwandtner, and Silvia Miksch are with the Vienna University of Technology, Austria. E-mail: {davide.ceneda,theresia.gschwandtner,silvia.miksch}@tuwien.ac.at.*
- *Thorsten May is with the Fraunhofer IGD, Darmstadt, Germany. E-mail: thorsten.may@igd.fraunhofer.de.*
- *Marc Streit is with the Johannes Kepler University, Linz, Austria. E-mail: marc.streit@jku.at.*
- *Hans-Jörg Schulz and Christian Tominski are with the University of Rostock, Germany. E-mail: hjschulz@informatik.uni-rostock.de; ct@informatik.uni-rostock.de.*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: 10.1109/TVCG.2016.2598468

The problem is that users, which are typically experts in their domain, but novices when it comes to VA, could be easily overwhelmed. Which method to use, how to set parameters, or how to get from one part of the data to another? Particularly when visual analysis methods are not applied on a regular basis, but only occasionally, such questions are not easily answered, a fact that hinders the effective use of VA in practice. What is needed are solutions that guide the user during data analysis and exploration. We see appropriate guidance as a key factor for significant improvements of the overall quality of data-intensive analytical work. In this context, the study and the development of tools for and models of guidance in VA is an important research topic.

While there are already a few approaches that offer guidance to users, there is only limited knowledge about the general mechanisms and underlying structures of guidance. Therefore, the goal of this paper is to contribute to a conceptual characterization of guidance. In Sect. 2, we study the design space of guidance and develop a general model of guidance in the context of VA. We build upon the initial characterization of guidance by Schulz et al. [42] and revise it with respect to the knowledge gap of users, the input and the output of a guidance generation process, as well as the degree to which guidance is provided (see Fig. 1). Van Wijk's [47] model of visualization serves as the basis for the development of a first model of guided VA. Our new model includes the fundamental building blocks of guidance and attaches them properly to the classic components of VA.

Sect. 3 bridges the gap between our conceptual considerations and guidance in practice. The individual dimensions and categories of the design space will be used to structure a review of existing approaches, which offer guidance in diverse ways. Selected examples from our own previous work will be described in more detail. In Sect. 4, we focus on open research questions related to guidance. With this we hope to stimulate the development of effective guidance approaches and systems in the future.

In summary, the key research contributions of this work are (1) a characterization of guidance in VA, (2) a conceptual model of guided VA, (3) a review of guidance approaches, and (4) a compilation of open research challenges.

2 Guidance: Terminology and General Concepts

In this section, we characterize the main aspects of *guidance*. In order to make this concept clear, we will first take a look at an illustrating example that deliberately leaves out any VA-specific aspects.

2.1 An Illustrating Non-VA Example

We imagine a smart car, supporting its driver in the journey to a destination. If the driver is confident about how to get there, he or she will drive the car, while the car provides guidance by showing the names of the traversed streets, highlighting the position of stops or traffic lights, and streaming the weather conditions for the current day. If the driver does not know how to reach the destination, the car could provide a higher degree of guidance by displaying turn-by-turn navigation instructions. These could also include alternative paths fulfilling certain constraints (e.g., avoid traffic jams or refuel required). Finally, in an advanced scenario, it is the car that drives autonomously to the destination, taking on each decision, changing paths if needed, but leaving the driver the freedom of taking over the steering wheel to deviate from the route or act in unexpected situations.

With this car example we sketch three different scenarios in which a system offers support to a human operator. By exploiting information derived from different sources and sensors, the system provides the driver with different degrees of assistance in order to address different needs: driving autonomously, searching for routes, and displaying additional information.

The example already hints at some of the important questions related to guidance. What are the needs of the human? How much guidance is provided by the system, and how is it conveyed to the driver? Based on what information is the guidance generated? In the next paragraphs, we will look at these questions in detail and through the lens of VA.

2.2 Definition of Guidance

Guidance is a broad term with much room for interpretation. To arrive at a crisp definition of guidance in VA, it makes sense to first review how the term is used in general and in related areas. Naturally, definitions provided in dictionaries are generic. According to two dictionaries, guidance can be defined as "advice or information aimed at resolving a problem or difficulty" [35] or "the act or process of guiding someone or something" [36]. These definitions are quite interesting, because they highlight guidance as a *process* aiming at solving a *problem*.

Another perspective of guidance is given in the field of human-computer interaction. Engels [14] outlines the main dimensions of guidance: the 'What', clarifying the problem, composed by an initial state and a goal state, and the 'How', aimed at solving the discrepancies between the two states by decomposing the main problem in a sequence of sub-problems that are easier to solve. Instead of focusing on the process itself, Smith and Mosier [44] emphasize the importance of interactivity and the visual nature of guidance defined as a "pervasive and integral part of interface design that contributes significantly to effective system operation". They also include guidance in their guidelines on visual interface design. The importance of guidance is also underlined by Dix et al. [13]. Since each analysis system might be used by different kinds of users, it is inevitable that not everyone will understand it. This is where guidance is essential, in the sense of *knowing where you are or what will happen*. Guidance has to be unobtrusive to the user, and adaptive to the particular context, as the type of assistance a user requires varies and depends on many factors.

In the visualization literature, one can find several notions that are similar or related to guidance, including *recommendations*, *incentives*, or *assistance*. Schulz et al. [42] group these different notions under the common term *guidance*. In their thinking, guidance refers to methods that have the goal of providing dynamic support to users, for example, when exploring data or when finding the best visual mapping for presenting analysis results. In addition to that, they also consider guidance in terms of suggesting a suitable domain expert and an appropriate computational infrastructure to carry out particular tasks.

From the diverse interpretations of guidance in various fields, we derive a definition of guidance in the context of VA:

> Guidance is a computer-assisted *process* that aims to actively resolve a *knowledge gap* encountered by users during an *interactive* visual analytics session.

According to this definition, guidance is a dynamic process that aims to support users in a particular task. In general, any task can be decomposed into a series of actions or decisions that lead to a desired result. Guidance provides support for at least one of these actions in situations where a user is unable to identify, judge, or execute the action. Our definition also includes cases where the desired result is not known in advance, and thus, the actual task must be derived from previous actions. Yet, we do not consider guidance to take over the reasoning part. For example, guidance is not supposed to retrospectively *explain* what is shown in visual data representations and how or why it came about. Instead, guidance provides prospective assistance so that users can make sense of the data on their own.

It is important to note that our definition focuses on the human perspective of guidance in that the system is guiding the human user [22]. There is also the notion of human users guiding algorithms to improve analysis results, but this is not what we are addressing here. This will become clearer in the next paragraphs, where we sketch a model of guided VA.

2.3 Conceptual Model of Guidance

As a starting point for a first model of guidance in the context of VA, we use van Wijk's [47] model of visualization. We make a slight modification though in that we replace the term *visualization* by *visual analytics*. This makes clear that our model covers both visual and analytical methods. The model is shown in gray in Fig. 2. Boxes represent artifacts, such as data or images, while circles represent functions that process some input and generate some output. Visual and analytical means (V) transform data [D] into images [I] based on some specifications [S]. The images are then perceived (P) to generate some knowledge [K]. Based on their accumulated knowledge, users can interactively explore (E) the data by adjusting the specifications (e.g., choose a different clustering algorithm or change the perspective on the data). As such, van Wijk's model effectively conveys the iterative and dynamic nature of knowledge generation mediated through VA. This makes it perfectly suited to be expanded to a model of guided VA.

We attach new guidance-related components to the model, shown in blue in Fig. 2. A central position is taken by the guidance generation process (G). It is hooked up first and foremost with the user's knowledge [K]. The reason is that before we can take any measures of guidance, we need to know what the particular problem of the user is. Similar to the worldview gap [2], we coin the term *knowledge gap* to capture the actual deficit that hinders continuation of the data analysis. The guidance generation process (G) is further connected to sources of information based on which guidance can be generated. These sources include the original data [D], visualization images [I], interaction history or provenance [H], and domain conventions or models [D]. Taken together, these components represent the *input* to the guidance generation process.

On the *output* side, results of a guidance generation process can be delivered in various ways. Fig. 2 illustrates three different scenarios. Orienting provides basic guidance through visual cues [C]. Directing offers useful options or alternatives [O] that the user may or may not choose to follow. Prescribing directly operates on the specification [S] in order to automatically generate suitable visual results.

The main goal of guidance is to create and maintain an environment in which users are able to make progress and perform their tasks effectively. This dynamic progressive procedure is well expressed by the knowledge change (dK/dt) occurring as a consequence of the

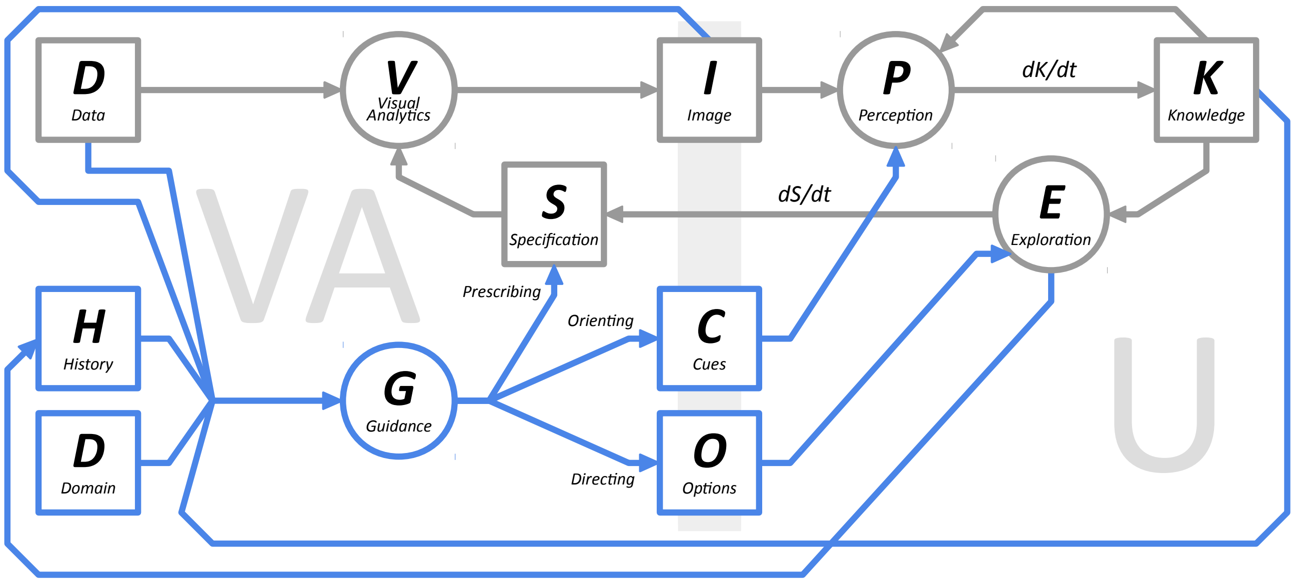


Fig. 2: Components of guidance (in blue) attached to van Wijk's [47] model (in gray). Aspects of guided VA are shown to the left, while user aspects (U) are on the right. Guidance considers the user's knowledge (or lack thereof) and may build upon various inputs, including data, interaction history, domain conventions, and visualization images. Different degrees of guidance are possible. Orienting uses visual cues to enhance perception. Directing supports exploration by providing alternative options. Prescribing directly operates on the specification. Guidance positively affects the user's knowledge in a dynamic process that eventually converges to zero knowledge gap.

guided visual analysis and the interactive adjustment (dS/dt) of the specification. A critical concern is that knowledge is acquired through perception and cognition (P). So the leverage point of guidance is to facilitate perception and cognition at different *degrees*, for example, by showing visual cues alongside the visualization, by offering options that, if chosen, lead to an improved visualization, or by taking over control and circumventing progress-hindering obstacles automatically.

In summary, we can identify three main characteristics of guidance: (1) the reasons why guidance is needed, i.e., the knowledge gap, (2) the inputs that are used to provide guidance as well as the output, and how the output is conveyed to the user, and (3) the expressed guidance degree. In the following, we will describe these aspects in detail.

2.3.1 Knowledge Gap

The knowledge gap pertains to the question: *What does the user need to know to make progress?* There are many different pieces of information that the user may need to know before progress can be made. It could be that a suitable color map has to be chosen before a certain data characteristic becomes visible. Or it may be necessary to visit different parts of the data before high-level relations can be discerned.

While a knowledge gap can come in myriad ways, there are two distinct *types* of knowledge gaps:

Target unknown means the user does not know the desired result. For example, the analyst has no idea about the clustering outcome to be generated.

Path unknown means that the user does not know how to reach the desired result. For example, given some ground truth, the analyst does not know which algorithm to choose and how to parametrize it, in order to extract the ground truth.

Fig. 1 illustrates the axis of known and unknown target and path that characterize the knowledge gap. Another perspective on the knowledge gap is the *domain* to which it pertains. There are five domains that are particularly relevant in VA:

Data: The user needs guidance in terms of data subsets or features. Guidance could (semi-)automatically identify such subsets or features based on some kind of "interestingness" definition, such as degree-of-interest functions or recommender systems.

Tasks: The user needs help in structuring a goal into a series of tasks that solve the goal. This is a high-level gap that guidance could narrow by hinting at what to do next. It is independent of the actual choice of VA methods to be used.

VA Methods: The user needs help with the available visual, analytical, and interactive methods. Guidance in this space could suggest suitable visualization techniques or algorithm parametrizations. This also relates to enhancements by means of providing additional information about VA methods.

Users: It is unclear who should carry out a task. When analysts work collaboratively, guidance could provide advice as to who would be a suitable expert to work on a specific task. This avoids situations where users are assigned to tasks that do not match their expertise.

Infrastructure: The user is unsure which infrastructure to employ. Guidance in this case means recommending hardware (e.g., display wall or touch-enable surface) and software (e.g., analytical mining tools or interactive exploratory tools).

Users may or may not be aware of the gap. It can very well be that a user does not even know that a certain procedure has to be performed before useful analytical results can be generated. This makes capturing the knowledge gap difficult. If users are aware of it, they can actively make it known to the system. If not, the system has to infer the knowledge gap, for example, by detecting deviations from domain conventions or long dwell times during exploration.

2.3.2 Input and Output

The input is concerned with the question: *What is the basis for generating the guidance?* When we look at the output of the guidance generation process, we are facing two questions: *What is the answer to the user's problem and how is the answer presented?*

The inputs are the foundations upon which guidance is generated. In the context of VA systems, we identified the following useful sources of information.

Data includes all kinds of information readily available or derivable from the data to be analyzed. Concrete examples are raw data, statistical properties of the data, data topology, or meta-data.

Domain Knowledge refers to information that originates from the application domain. This could be expert systems, domain models, workflows, or conventions.

Visualization Images include the visual data representations and information about mapping parameters. They can be useful for understanding what the user is actually seeing.

User Knowledge is about information that users input to the system, including annotations or degree of interest (DOI) functions, or information that the system can infer from the user.

History relates to keeping track of interactive changes. This includes logging interaction steps, employed algorithms, applied parameterizations, or visited parts of the data.

Concerning the output of the guidance generation process, there are two aspects to be considered: finding of a suitable *answer* and using appropriate *means* to convey the answer to the user.

Answer: Conceptually, finding the answer boils down to developing a function that takes the knowledge gap plus additional input and computes a suitable result.

$$guidance(gap, input) \rightarrow answer$$

This definition is abstract and broad enough to consider many different situations. Iterations of the function converge to the goal of zero knowledge gap, where each iteration conveys a variable amount of knowledge to the user, depending on the user's expertise and perceptual and cognitive abilities. In this sense, guidance is an active process and the user is included in the loop.

We distinguish *direct* from *indirect* answers. Usually, the knowledge gap should be answered directly. For example, if a user has a problem in finding a suitable value for a clustering parameter, the guidance generation process should provide promising candidates. On the other hand, guidance could provide indirect answers. Staying with the same example, the guidance could hint at interesting structures in the data, whose analysis (note the indirection) may help the user fine-tune the clustering parameter.

Means: Once computed, the answer has to be communicated to the user. This is a critical step. The goal is to induce an impulse in the user so as to enhance perception or to trigger exploratory actions. It is typical in VA settings that the answer is presented visually. This could mean adjusting the visualization mapping, providing visual enhancements, or including additional user interface elements. Yet, we do not consider the means to be limited exclusively to the visual channel. Depending on the context in which guidance is used, answers can be provided by exploiting non-visual channels as well, including sounds or tactile feedback.

2.3.3 Guidance Degree

The guidance degree is about the question: *How much guidance is provided?* For the car example mentioned earlier, we already saw that guidance can be provided at different levels. The same holds true for guidance in VA. The *guidance degree* specifies the extent to which guidance is required and actually provided. The guidance degree is not static, but varies over time as tasks, data, and procedures change through the course of a VA session. This enables guidance to be fine-tuned to the requirements at hand. For example, if a user gets lost during data exploration, the guidance degree should be increased. If the user feels too restricted by the system-prescribed course, the guidance degree should be decreased.

The two extremes of the guidance degree are *no guidance* (no support given to the user) and *fully automated* (no options for the user to intervene). These are, however, only of theoretical relevance. In practice, the guidance degree is in between these extremes, with three characteristic scenarios being particularly interesting to look at:

Orienting: Providing merely orientation is at the low end of the guidance degree. The main goal is to build or maintain the user's mental map. Orienting in VA typically involves adopting the map metaphor for an abstract domain. Such a map may contain potential targets and paths as well as relations among them. Providing visual cues hinting at these targets and paths are a common strategy for implementing orientation. Visual overview technique may provide some kind of orientation as well.

Directing: Directing represents a medium degree of guidance. In contrast to orienting, directing approaches emphasize a certain preference for a future course of action. The system presents the user with a set of alternative options to produce the desired result or a set of similar results. The suggestions may differ in terms of quality and costs for different paths leading to the same result or, in terms of interest for paths, leading to similar or new results. Directing can benefit from preview techniques that help users make informed decisions for one or the other option.

Prescribing: With prescribing we reach a higher degree of guidance. In contrast to directing, prescribing approaches make decisions on steps to be taken on their own. Prescribing implements a largely automated process, which proceeds towards a specified target. Such a process may cover any (sub-)task of analysis regardless of its scope. In the context of VA, it is important to visually present the intermediate steps of the process and the decisions that lead from one step to the next. In a sense, this degree of guidance can be compared to an interactive presentation. A user may interrupt the presentation and ask for details, or rewind/reverse it to revisit a nugget of knowledge that has been found earlier. Depending on the degree of automation, the user can recover control for a while and nudge the presentation to another path or even another target.

With these three scenarios we have completed sketching the key characteristics behind guidance. In the next section, we will use the developed characterization to structure a broader review of existing guidance approaches in the context of VA.

3 A Review of Guidance in Visual Analytics

There is no single comprehensive guidance approach for VA that covers all aspects that we discussed in the previous section. Yet, instantiations of specific aspects can be found in existing work. In this section, we apply our characterization to a selection of examples to showcase the state of the art and to show possible connections between complementing approaches.

3.1 Knowledge Gap

3.1.1 Type

The following examples illustrate the difference between guidance approaches allowing the user to find and specify solutions, and guidance approaches that allow the user to pursue the path towards a solution.

Target Unknown The *target* refers to a solution to a specific problem, such as a useful visualization. Usually such a solution is not purely deterministic, but instead is defined in guided interaction with the user. For instance, Fujishiro et al. devised Gadget [15], a knowledge-based tool that builds upon Wehrend's task taxonomy [50] with the aim of suggesting a set of possible goal-oriented visualizations. Another approach is BOZ by Casner [10], which models tasks as a set of logic rules and designs a possible equivalent perceptive process to provide the user a set of views. These views aim to support the user's perceptual capabilities and improve the user's performance. Both approaches provide support in choosing the correct target, in these cases a visualization. The users of automated techniques face similar problems. Choosing appropriate techniques for an analytical task or selecting their parameters are cases of unknown targets. As one of many examples, Krause et al. [25] developed a tool to rank data features for modeling, offering guidance in the feature selection process. In this case the target is the set of most useful features.

Path Unknown The next two approaches address the problem of finding sequences of actions to achieve a goal, be it the creation of a view or the application of filters to a dataset. Willet et al. [51] developed scented widgets, a technique that offers guidance in the data domain, to help users in completing a series of data transformation steps. These widgets are interactive elements in a graphical user interface that incorporate information about other users' activity. The hints provided by scented widgets level possible knowledge gaps and lead inexperienced users to significant results. The visual pre-processing by Bernard et al. [6] offers guidance in composing a sequence of steps for time series transformation. The effects of each step are demonstrated by input-output comparison of time series samples suggested by the system.

3.1.2 Domain

The guidance domain captures the subject matters with respect to which a knowledge gap can manifest. Most of the existing literature is concerned with guiding towards data of interest and suitable VA methods. Yet, the following approaches will illustrate how versatile the guidance domain can be beyond data and VA methods.

Data Finding data that are worthwhile to investigate in a large dataset is a known challenge in VA research. One of the most prominent ways of assisting this task is by capturing what makes a data item interesting to the user in a so-called *degree-of-interest function* and recommending those data items with high interest values to the user [16]. Aspects that factor into such a quantitative notion of interestingness are, for example, special data characteristics (e.g., uniqueness, extreme properties), novelty (e.g., whether a data item has been looked at before), or visual saliency (e.g., whether a data point is visible or overplotted). To infer automatically what parts of the data might interest the user is subject of the area of *user profiling* and in particular *preference elicitation* [19].

Tasks Given some data of interest, it is not necessarily obvious what to do with it. There is a large variety of potentially relevant tasks to be executed next [41]. Step-by-step methodologies or analytical workflows that have been found to be generally good approaches in a certain domain can help in such cases to suggest promising analytical tasks. An example for such a methodology is given by Perer et al. [37] for network analysis. Using the knowledge of domain experts, their methodology specifies intermediate tasks that a user needs to carry out in order to yield a correct end result. While the tasks are fixed, the user has the freedom of completing them in any order.

VA Methods This domain can be used to provide more concrete guidance in terms of "how to do a task?" than just suggesting a task. Offering guidance on VA methods means to point out concrete tools or algorithms to use. The assistant *VizAssist* [8] provides such guidance by matching the data to be analyzed to suitable visual methods. Increasingly better suggestions are derived by evolving the matchings through an interactive genetic algorithm that takes user feedback into account.

Users In relation to collaborative VA, Heer and Agrawala [20] asked already in 2008: "Can automated techniques be used to help allocate effort?" They envisioned that tapping into user profiles, logs of prior analysis contributions, and social networks could help to automatically determine suitable collaborators with the expertise necessary for a particular task. This actually relates closely to the field of *expert finding* [32], for which already some visual and analytical tools exist [31]. In the field of VA, these methods are not yet picked up on.

Infrastructure As computing power and display spaces become increasingly ubiquitous these days, it is no longer evident on which device to perform which task. Radloff et al. [39] present a framework for smart view management, that takes views, available display spaces, and analytical tasks into account to suggest favorable mappings onto available displays. In essence, it computes for each possible view-display mapping a *view quality score* that is weighted by the importance of the view for the task at hand. Thus, the framework suggests view configurations that maximize the sum of these weighted scores.

Each of the guidance domains above represents a research challenge in its own right and most guidance approaches address exclusively one of them. Yet, in principle it is possible to combine them, as it is for example done by Streit et al. [45]. Their approach captures multiple domains in a set of interlinked models that contain information on datasets, tasks and workflows, preferred visual and analytical methods, as well as different user expertise needed for those datasets. These models can then be leveraged to extract analytical recommendations from them.

3.2 Input and Output

3.2.1 Input

Inputs are the sources of information that are used to generate guidance. Most approaches require a combination of sources to offer a useful solution. Our examples are categorized according to their primary source.

Data Gratzl et al. created Domino [18], a general technique for tabular data that permits the user to create, explore and extract heterogeneous data subsets and show their relationships by visually connecting them. Visual cues indicate compatible views, with respect to data properties. Lex et al. designed StratomeX [28], which aims to help scientists in identifying cancer subtypes. The tool derives and highlights cancer subtype relationships across different datasets. In both these examples, data properties and connections among data entities are exploited and mapped to visual elements, such as lines or ribbons, in order to make them clearly visible to the user and to provide a defined context to enhance the user's orientation and awareness.

Domain Knowledge Guidance can also be generated based on domain related knowledge: task knowledge, workflows, and conventions. The work by Streit et al. [45] presents a step-by-step process for the analysis of heterogeneous data. The process aims to satisfy both experienced and inexperienced users improving orientation and analysis completeness by using tasks knowledge and providing the user a clear sequence of steps to reach a result. In general, there are many approaches that use domain knowledge to generate guidance. Some of those we have already discussed in previous sections of this paper [10, 15, 37].

Visualization Images This category focuses on guidance systems that exploit information derived from views, mappings, and visual elements. One example of taking visual features as input to generate user guidance is given by Wang et al. [48]. They devised a guidance approach in the field of graph drawing. It provides guidance by calculating an index about the ambiguity of the graph drawing (e.g., edge crossings or insufficient distances among nodes) and highlighting problematic graph regions. This approach considers the visualization at hand to guide the user on which areas to investigate further in order to uncover cluttered parts of the represented network.

User Knowledge User feedback, be it explicit (the user evaluates his/her experience directly) or implicit (the information is deduced from the user's actions and performances), is also a valuable input for generating user guidance. While implicitly derived feedback avoids cumbersome feedback collection and does not interfere with the user's workflow, it may be subject to errors caused by misinterpreting the user's activities [34]. Mouse events, like clicks or hovering over specific regions of the display, are a source of implicit information about the user's preferences and interests. It could be used, for instance, to steer a document retrieval operation or the search for a specific product in an e-commerce website [21]. Gotz and Wen [17] present a comprehensive example of user and task based guidance. The interaction log of the user is matched with a set of interaction patterns derived from previous user behaviors. These patterns are used to identify the implicit task, which in turn is used to adapt the visualization.

History Another possible input for generating guidance is information derived from the user's past exploration process. Kreuseler et al. [26] introduced a history management unit to be included in a visual data mining framework. This tool represents the historical sequence of operations as a tree, with undo and redo functionality. A similar history visualization is presented by Derthick and Roth [12]. These two solutions foster orientation in that they sketch the paths that have already been explored, which allows users to orient themselves with respect to previous analysis actions. Shrinivasan et al. [43] proposed a tool that consists of three views, of which one is intended to show the analytical process history, one represents the findings, and the last one shows the dataset. These views enable the user to build a context that can help justify or prove a result or finding.

3.2.2 Output

The output of the guidance generation process is composed by the answer to the user's knowledge gap and by its (visual) representation. It may happen that the output of the guidance generation process does not fully satisfy the user requirements at first, however the output can be seen as an iterative function that converges towards zero knowledge gap. With each iteration the user acquires more knowledge with respect to the problem at hand.

Answer Although the answer corresponds to a user need, which is a direct consequence of the knowledge gap, we consider also the case in which guidance, and thus the answer, is not directly offered to the user but is provided indirectly.

Direct: The answer is given on the same domain as the knowledge gap. The approach by Perer et al. [37] is related to the question: *Which are the steps to reach the result?* As the system provides the user with a list of steps to complete the task, it provides the answer in a direct manner. The approach by May et al. [30] deals with the lack of knowledge in finding interesting graph regions as well as the shortest path to reach them: the proposed solution guides the user by showing interesting regions represented by signposts, and indicating the shortest path leading to them.

Indirect: Approaches falling into this category include [18, 28]. In these examples, the knowledge gap coincides with: *The user does not know which is the best way to visualize and compare subsets*, and *The user would like to mix different data sources*. However, the systems do not guide the user directly to results, but instead take care of the visualization of subsets or relationships among them. The approaches offer orientation by providing a meaningful context in which the user performs the task to gain insights. In other approaches that fall into this category [12, 26, 43] the knowledge gap relates generically to gaining insights. However, the user is just supported in the trial-and-error process by making explicit the history of actions.

Means Once an answer is computed, it has to be communicated to the user. In VA the output of a guidance generation process is usually provided by adding, changing, or removing elements from the current view, or by providing interaction facilities to support the exploration process. In Stack'n'Flip [45], the authors propose a visualization in which the sequence of steps needed to perform a task (i.e., the answer to a user need) is visually shown and added to the view: the path to follow is added below the main view together with the needed datasets. Alternative paths are of different color, while possibly related paths are highlighted. Jankun-Kelly and Ma [23] present an approach to guide the selection of parameter combinations in huge parameter spaces. The key idea is to present the user with a stack of two-dimensional spreadsheets showing all possible combinations of dimensions. Dedicated interaction techniques support the navigation in the parameter space. The user can then easily explore suitable parameter combinations for the problem at hand. Similarly, Lehmann et al. [27] propose a method to generate pictograms for communicating specific properties of data distributions in multidimensional visualizations, in order to ease judgement of these properties. Some examples of interaction facilities that follow from a guidance generations process include Kreuseler et al.'s [26] or Derthick et al.'s [12] history mechanisms which support undoing and redoing of actions. Scented widgets [51] are interactive elements of a user interface enhanced with visual suggestions: the actions performed by other users are visually added and summarized in form of hints around the control, while the relevancy of each option is underlined by using different color schemes.

3.3 Guidance Degree

Another important aspect of guidance methods is the degree of assistance provided, which should meet the user's needs. It is a continuous spectrum that spreads from orienting to prescribing.

3.3.1 Orienting

Support for orientation is closely related to the goal of building and preserving a user's *mental map*. A mental map is a spatial representation of a real space, or of abstract relations in possibly any topic of interest. The relevance of a mental map has been recognized in various studies in the field of graph drawing [3, 38]. Like a real map, it serves fundamental orienting tasks like path-finding, self-location, or exploration.

A mental map for VA typically spatializes abstract relations. We present two groups of examples that operate in two different domains of the knowledge gap. Approaches in the first group primarily offer orientation in the *data* domain. These approaches aim at mapping relations between data subsets, patterns, attributes or models. Gratzl et al. [18] and Lex et al. [28] help users understand these relations by showing the connections between different parts of the data. Some of the relations may be known beforehand, others may be introduced during analysis. With a similar goal, Yang et al's. [54] approach offers orientation in the 'pattern space'. It generates a map of patterns found during an entire session. The patterns are arranged according to their similarity, regardless of how and when the patterns actually have been defined.
Approaches in the second group primarily offer orientation in the *task* domain. These approaches aim at spatializing the series of tasks in the analytical process. This may include methods or intermediate results as well. Kreuseler et al. [26] sustain user's orientation by making explicit the history of actions, thus, providing guidance in trial-and-error systems. Shrinivasan et al. [43] subdivide the analysis process by assigning different views to the history of actions, datasets, and findings, with the aim of supporting the exploration. Finally, approaches like the one proposed by Streit et al. [45], provide orientation but as a part of a broader guidance support: in this case data properties, relationships between datasets and predefined domain-specific workflows are exploited to provide assistance.

3.3.2 Directing

Directing approaches offer a ranking or preselection of alternatives, which can be inspected and finally selected by the user. Koop et al. [24] propose an approach for the creation and completion of visualization pipelines. The knowledge source is a database of previously created visualizations. While the user creates a pipeline, the user is offered suggestions for the most frequent completions. VizAssist [8] and Voyager [53] are recent examples for guiding the choice of visualizations in the context of an analytical process. Both approaches focus on guiding the selection of data and the mapping, rather than on guiding through the visualization design-space. Both use expert knowledge, automatically generated rankings about the data, and user intentions as guidance input. Remarkably, in VizAssist, user intentions are defined explicitly from a catalogue. In Voyager, implicit user intentions are defined incrementally via variable selection.

The guided improvement of visualizations can be complemented by techniques for improving analytical results as generated by different algorithms under different parametrizations. *Directing* approaches in this category display multiple, selectable parameter settings in relation to the quality of results. Bernstein et al. [7] propose an approach for the assessment of classification models and modelers. Infuse by Krause et al. [25] combine the assessment of classifier and feature selection methods. In terms of our characterization, these examples aim at bridging the knowledge gap in the domain of VA methods.

3.3.3 Prescribing

While techniques that provide directions allow users to follow or ignore them, prescriptive guidance approaches purposefully limit user influence to traversing a fixed path of analysis. The reasons to do so can be manifold, for example, to reduce the learning curve for casual users by providing them with a simplified analysis experience [5], to streamline the analysis process in potentially "distraction-rich" datasets [1], or to have the analyst stick to an agreed upon standard operating procedure or best practice for better comparability or reproducibility of the results [45].

On a user interface level, this guidance strategy is epitomized by the *wizard interface*. It leads users through a complex task by breaking it into a sequence of smaller tasks that can be carried out step-by-step. Streit et al. [45] show a modern incarnation of such a wizard for visual analysis that departs from the classic modal dialog featuring two buttons to navigate back and forth among the subtasks. Their Stack'n'Flip interface, collects data visualizations that were already explored on one side, those that still need to be explored on the other side, and the one that is currently being explored in the middle of the screen. A linked visualization of the workflow serves as a navigation aid to go back and forth through this stack of visualizations. While still allowing deviations from the workflow, this interface discourages them and shows analysts how to get back on track.

On the view level, the prescriptive guidance strategy is embodied by the concept of providing a "tour" through the data. This idea originated from Asimov's work on the *grand tour* in high-dimensional data spaces [4]. At its core, it is an animation of different 2-dimensional projections of a multivariate dataset in an attempt to show the data from all possible angles. This idea has since been applied to other types of data, as well. For example, Yu et al. [55] present a mechanism that automatically constructs such an animated tour from events in time-varying data, whereas Wohlfart and Hauser [52] developed an approach that creates a guided and interactive visual story for volume data. While the story is completely defined by the system, the user is left the freedom of asking for details as well as interacting with the story playback. More abstractly, Dennis and Healey [9] provide a framework for data spaces in general, called *assisted navigation*. It can be used to generate tours that span certain elements of interest in data space as well as areas of interest in view space.

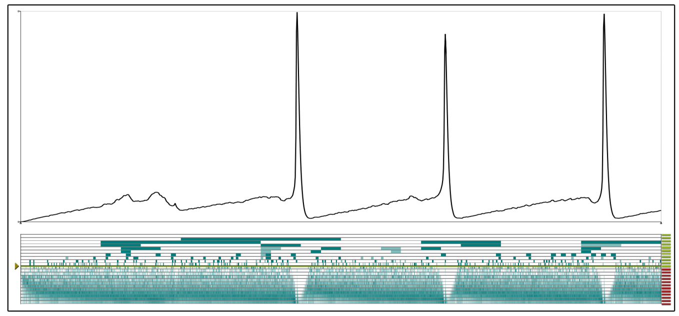
3.4 A Detailed Look at Selected Examples

In the previous paragraphs, we provided an exemplification of each single characteristics of guidance in VA. Next, we will be looking at three approaches in detail. To the best of our knowledge, no approach covers the whole guidance spectrum. Yet, the following examples highlight the most relevant factors when characterizing guidance.

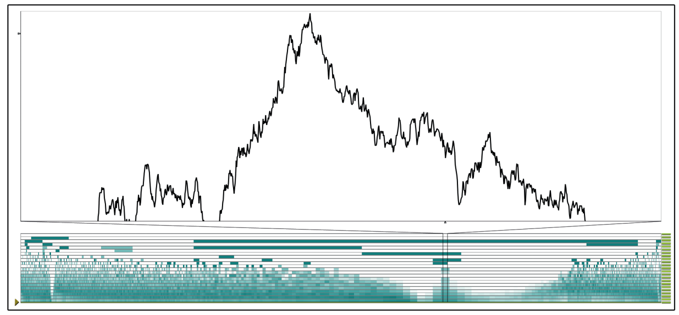
3.4.1 Example 1: Heterogeneity-Based Guidance

Luboschik et al. [29] facilitate the exploration of multiscale data. The approach points the analyst to scales and regions within the data (*unknown targets*) that exhibit behavior of interest without the need for an exhaustive search. The main idea is to take the most fine-grained *data* as a *guidance input* and to step-wise aggregate it into more coarse-grained data. Pairs of subsequent data scales can then be compared by various metrics, detecting data features that were observable in the more detailed scale, but can no longer be found in the less-detailed aggregated scale. In other words, subsequent scales exhibit heterogeneous behavior. This information is then communicated to the user by *means of visual cues*, in this case colored heterogeneity bands that provide *orientation* towards regions that are worthwhile to zoom into. This way, the analyst is given a *direct answer* to the question where deviating behavior from the currently shown will emerge, while at the same time not having to bother with investigating other parts of the data where no such deviation occurs.

Fig. 3a shows an example of this approach, where a lineplot of millions of data points (top) is enriched with a display of multiscale heterogeneity bands (bottom) that measure how well slope changes are preserved between subsequent scales. The heterogeneity bands show three valleys and within them, very thin, suspicious peaks exactly



(a) A lineplot (top) enriched with multiscale heterogeneity bands (bottom).



(b) A zoomed view of one of the spikes.

Fig. 3: Orientation by means of visual cues [29]. (a) The lineplot shows clear spikes among millions of data points. The heterogeneity bands below the plot suggest that there is more to these spikes hidden at higher levels of granularity. (b) Zooming in on one of the spikes in confirms this assumption.

at those points where the lineplot is at a maximum. Guided by this indicator of more nuanced behavior at these points, the analyst zooms into one of these instances in Fig. 3b. One can immediately observe that the maximum is far from being as clearcut as the overview in Fig. 3a suggested. Instead of a distinct tipping point, upward and downward movements are at a constant struggle against each other, until the latter gets the upper hand and reverses the strong upward trend. Without guidance, this interesting behavior of the data at a more detailed scale would have gone unnoticed or only be found by pure chance.

3.4.2 Example 2: Signposts for Navigation in Large Graphs

May et al. [30] support the orientation in large graphs by using glyphs representing signposts as shown in Fig. 4. The sign posts are inspired by their real-world counterparts. Only a small subgraph is shown at any time. Orientation is supported by pointing to labeled regions of the graph outside the visible area. The signposts are attached to outbound edges connecting the focal area to the invisible regions along the shortest paths. The signs that are actually shown in the view are selected by the relative importance of regions in terms of distance, region size and overall graph coverage. Moving the visible subgraph triggers a recalculation of the relative importance of regions, and thus the selection of their signs.

In terms of the characterization of guidance, the signposts approach is a technique for orientation with the guidance domain being the graph data themselves. The primary knowledge gap addressed by the approach is literally an *unknown path*. A user will reach any region of interest by following breadcrumbs. Hence, the guidance output is a glyph, which indicates the beginning of the shortest path, and offers an affordance to short-cut movement directly to the target region. To associate a signpost to an intended target, a user requires meaningful names for any region given. The guidance input is based on interaction history and user knowledge. Firstly, the history of visited focal areas is maintained to assess region importance. Secondly, user-defined regions are stored as priority landmarks to ease revisiting.

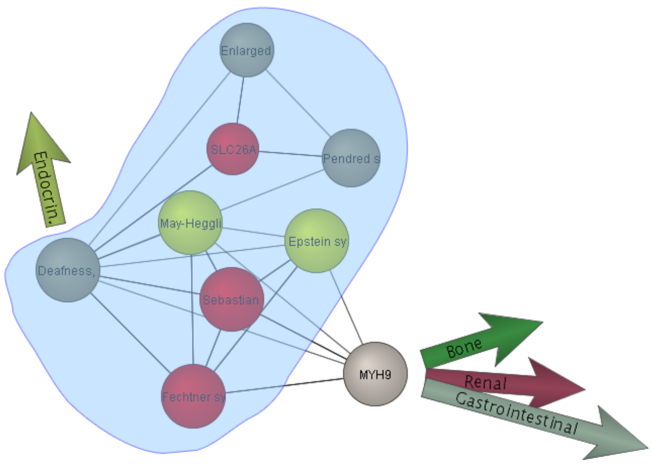


Fig. 4: Orientation via signposts [30]. Signposts connect a small, yet detailed focus region of a graph to the invisible "context". They label outbound edges that connect invisible regions along their shortest path.

3.4.3 Example 3: Model-Driven Guidance

In the work by Streit et al. [45], analysts are guided through an analysis session based on a predefined comprehensive model as depicted in Fig. 5a. The model, which is defined in an authoring process, consists of three stages: (1) a setup model, describing how heterogeneous datasets are connected and which visual and computational interfaces can operate on the datasets; (2) a domain model which defines domain specific tasks and their relation to the setup model; and (3) an analysis session model that defines a workflow as a sequence of tasks.

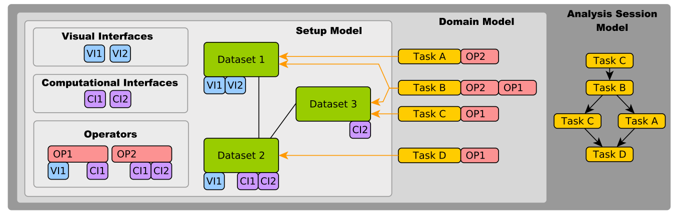
During an analysis, the setup model serves as a basis to orient the user during an analysis session in the domain of tasks and methods. Hence, the guidance degree is characterized by both orienting and directing. As the workflow is predefined, the path is known, while the target is unknown. The guidance input is covered by the three-stage model (data, visual and analytical interfaces, workflow, and domain specific tasks) together with the history of the analysis and further user input, such as user-defined thresholds. The guidance output is a tree-based meta-visualization that is used for both orienting and directing the analyst, as shown in the lower part of Figure 5b.

4 Discussion and Future Work

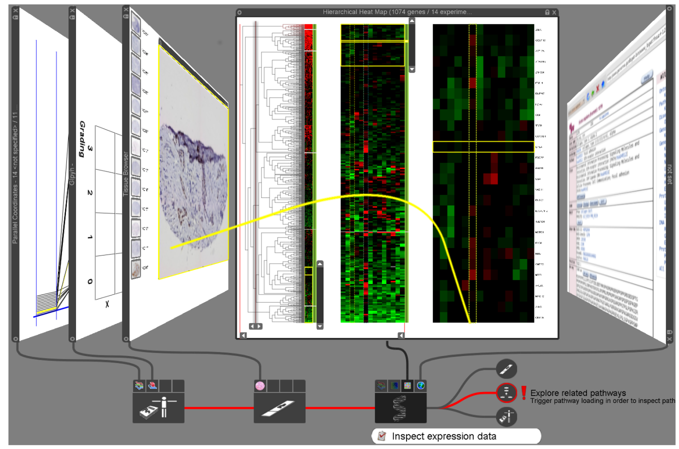
In the previous section, we have seen how existing guidance techniques can assist the user in various ways. The model, as introduced in this paper, is a first step to systematize the emerging field of guidance in VA. In this section, we identify open research questions and derive suggestions for future work on guidance.

Refining the model Our model explains the embedding of guidance in VA scenarios. It comprises the fundamental components of guidance and their interplay. This helps us understand how guidance works in principle. A sensible next step for the future is to refine the model to develop a better understanding of the internals of guidance. For example, the core function of guidance, i.e., the guidance generation process, largely remains a black box. The illustrating examples implement it in one way or the other. Yet, it remains to be studied if one can extract a general procedure of *how* guidance is actually generated. Such a procedure could then be used as a blueprint for developing new guidance techniques. A sensible refinement to our framework may come by known models. Sacha et al. [40] expanded the original VA pipeline to highlight the strong synergy between human and machines while generating new knowledge. In the same way it is possible to look at the guidance model to spot where and how it is possible to provide assistance both to the human and to the machine loop.

Similarly, our understanding of the knowledge gap remains limited. Most existing approaches either implicitly infer knowledge gaps



(a) Domain-specific three-stage model.



(b) Based on the model, stack'n'flip guides users through various analytical views.

Fig. 5: Model-driven guidance [45]. (a) A domain-specific model is defined in a three-stage process. (b) The model is then utilized to support users during the data analysis.

a-priori from overplotting and other ambiguities in the visualization (What parts of the data are not visible to the user?) or a-posteriori from interaction histories (What parts of the data the use has not explored yet?). It remains an open challenge to do the same during an ongoing analysis. Simple heuristics, such as long idle time, can be used to automatically detect stalled analysis sessions. Such methods provide but simple indicators of the fact that guidance is needed. For well-balanced and effective guidance, the knowledge gap needs to be specified in greater detail. A promising starting point is to consider established models from human-computer interaction. In Norman's action cycle [33], the execution phase is associated with three layers of competence, *knowing why*, *knowing what*, and *knowing how*. All are needed for making progress in a human-in-the-loop analysis process. Distinguishing these layers will allow us to better attune guidance to the user's personal level of competence. To this end, a fundamental approach to identifying the knowledge gap during the analysis is needed. However, the back and forth between diverging processes (exploration) and converging processes (confirmation), which is typical for VA sessions, makes this a formidable research challenge.

Novel guidance approaches In the literature, there are a number of approaches that deal with guiding in selected aspects of VA. However, we did not find any guidance approach that covers the entire VA process. Here we see potential for future work on novel guidance techniques. New techniques could specifically address the lack of comprehensive guidance for the human-in-the-loop process and offer intertwined guidance on all phases of VA (e.g., how to transform data, modify calculations, and how to read and interact with the resulting visual representations). Just as we see a specialization of VA for specific data classes (e.g., multivariate data, graphs, text), we believe that it also makes sense to consider tailored guidance approaches. An example are guidance techniques for time-oriented data. The dimension of time has a rich structure and it is not always clear to the analyst which facet of time to focus on (e.g., linear time vs. cyclic time). Navigation in time is another aspect where guidance could assist the user in visiting those parts of the data that potentially lead to interesting findings.

When we look at existing techniques, the majority of them generates guidance based on the data (e.g., [18, 28]), past analytical actions (e.g., [12,26,43]), or planned future analytical actions (e.g., [10,37,45]), such as workflows, analysis protocols, or standard procedures from the application domain. Only a few techniques (e.g., [48]) consider the visual representations as input to generate guidance. What other inputs can be useful, emotions [11] for example? Another limitation is that current approaches typically consider only a single type of input. Particularly in the light of the different layers of competence as indicated before, there is a need to consider multiple sources of information. However, it is still an open question how various inputs can be combined in general.

On the output side of guidance, we have a similar situation: Most techniques provide only one degree of guidance: orienting, directing, or prescribing. Novel guidance approaches should support adaptively switching between guidance degrees in order to generate a richer experience. For example, if the user deviates often from the proposed route, orienting may be more suitable than directing or even prescribing. More research is needed to investigate mechanisms for triggering switches between degrees. What would be appropriate indicators (e.g., user input, situation monitoring) and suitable thresholds for automatic switching? Moreover, the guidance interface needs to be designed so as to make switches in the degree transparent to the user.

Regarding the human, existing approaches typically assume a single individual. Yet, VA is increasingly a collaborative effort of several analysts. So far, there are only very few approaches that offer guidance in collaborative scenarios. This is a largely open research question.

Evaluation of guidance Evaluating visualization techniques is notoriously challenging. VA with its mix of analytical, visual, and interactive methods is even harder to evaluate. On top of that, guidance adds considerably to the evaluation challenge. The tight coupling among the involved methods makes it difficult to set up controlled experiments. Already when investigating the visual embedding of guidance (what we refer to as *means*), a number of evaluation questions come to mind. For example, which means are appropriate for what tasks or which means are best suited for which degree of guidance?

Moreover, faster completion time and fewer errors alone might be insufficient to draw conclusions about the usefulness or utility of guidance approaches. An interesting alternative question is if guidance sends the user along worn-out paths or if it is able to suggest side tracks to allow for unexpected discoveries. One way to evaluate this is to simulate the use of guidance. To this end, one can pseudo-randomly select from the suggestions generated by guidance and mark the corresponding spot in the data or parameter space as visited. Useful guidance would lead to the relevant parts of the data or parameter space being gradually filled with marked spots.

Another suggestion to tackle the challenge of evaluation, is to consider self-reporting methods. Ideally, guidance would monitor the situations in which the user resorts to it and keep track of its use. This would allow for deriving conclusions about the utility of guidance depending on the different situations during visual data analysis. Moreover, the collected information can be used not only for evaluating guidance, but they could also serve to implement self-adapting or learning guidance. Certainly, this would require combining guidance with concepts known from artificial intelligence.

Guidance and guidelines With our work, we structure the space of guidance solutions. While guidance is to support the user in using VA tools, we have not considered *guidelines* that apply in the development phase of VA. Particularly with guidance for different data and different tasks, and maybe even for different users employing diverse infrastructures, it can become difficult to develop or choose an appropriate guidance technique for a given problem. Therefore, it is important to provide both guidance for users and guidelines for developers. By guidelines we mean established best practices that a developer can refer to when implementing VA approaches. Such guidelines could, for example, suggest how certain analytical situations are best supported with a certain degree of guidance. We see much potential for future research on guidelines enabling us to make the most of guided VA.

From guidance to mixed initiative visual analytics In this paper, we focused on guidance generated by the computer and provided to the user. Yet this thinking is limited in that it considers only one direction of guidance. Much of the potential of VA lies in the close cooperation of human and computer. To fully exploit this potential, it is necessary to include users assisting the computer in the guidance equation. The benefit of user interaction for complex problem solving has long been known [49]. Yet it remains challenging to integrate human and computer on equal footing to obtain VA solutions that are truly mixed initiative. To tackle this challenges, we first need to better understand the back and forth between computers guiding humans and humans assisting the computer.

5 Conclusion

In summary, our work contributes to a better understanding of guidance in VA. We defined guidance as a dynamic, iterative, and forward-oriented process that aims to help users in carrying out analytical work using VA methods. Guidance was further characterized along the knowledge gap of the user, the input and output of the guidance generation process, and the degree of guidance that is actually provided to users. We developed a first conceptual model of guided VA based on van Wijk's model of visualization. A structured review of existing approaches illustrates diverse ways of how guidance can be applied in the context of VA. Finally, we identified open research questions to be addressed by future work on guidance.

In conclusion, we established a basis for the comprehension and the development of assistive approaches that improve the insight generation process and ease the visual exploration and analysis of data.

Acknowledgments

We thank the participants of the Rostock Workshop on Emerging Topics in Visualization and Computer Graphics 2013 (WET-VCG) for valuable initial discussions on guidance. The presented research drew inspiration from the Dagstuhl Seminar 13352 on Interaction with Information for Visual Reasoning. This work was supported by the Centre for Visual Analytics Science and Technology CVAST, funded by the Austrian Federal Ministry of Science, Research, and Economy in the exceptional Laura Bassi Centres of Excellence initiative (#822746). Further support has been received from the State of Upper Austria under grant number (FFG #851460). Finally, we thank the anonymous reviewers for their helpful comments and valuable feedback.

References

[1] S. Alspaugh, A. Swigart, I. MacFarland, R. Katz, and M. Hearst. Rethinking the first look at data by framing it. Technical Report UCB/EECS-2015-208, University of California at Berkeley, 2015.

[2] R. Amar and J. Stasko. Knowledge precepts for design and evaluation of information visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):432–442, 2005. doi: 10.1109/TVCG.2005.63

[3] D. Archambault and H. C. Purchase. Mental map preservation helps user orientation in dynamic graphs. In *Graph Drawing*, pp. 475–486. Springer, 2013. doi: 10.1007/978-3-642-36763-2_42

[4] D. Asimov. The grand tour: A tool for viewing multidimensional data. *SIAM Journal on Scientific and Statistical Computing*, 6(1):128–143, 1985. doi: 10.1137/0906011

[5] G. Atemezing and R. Troncy. Towards a linked-data based visualization wizard. In *Proc. of Conference on Consuming Linked Data (COLD)*, pp. 1–12. ACM, 2014.

[6] J. Bernard, T. Ruppert, O. Goroll, T. May, and J. Kohlhammer. Visual-interactive preprocessing of time series data. In *Proc. of SIGRAD*, vol. 81, pp. 39–48. Linköping University Electronic Press, 2012.

[7] A. Bernstein, F. Provost, and S. Hill. Toward intelligent assistance for a data mining process: An ontology-based approach for cost-sensitive classification. *IEEE Transactions on Knowledge and Data Engineering*, 17(4):503–518, 2005. doi: 10.1109/TKDE.2005.67

[8] F. Bouali, A. Guettala, and G. Venturini. VizAssist: An interactive user assistant for visual data mining. *The Visual Computer*, pp. 1–17, 2015. doi: 10.1007/s00371-015-1132-9

[9] D. Brent and C. Healey. Assisted navigation for large information spaces. In *Proc. of IEEE Conference on Visualization*, pp. 419–426. IEEE, 2002. doi: 10.1109/VISUAL.2002.1183803

[10] S. Casner. Task-analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, 10(2):111–151, 1991. doi: 10.1145/108360.108361

[11] D. Cernea, A. Ebert, and A. Kerren. A study of emotion-triggered adaptation methods for interactive visualization. In *Proc. Workshop on User-Adaptive Visualization (WUAV)*, pp. 9–16. CEUR-WS.org, 2013.

[12] M. Derthick and S. Roth. Data exploration across temporal contexts. In *Proc. of Conference on Intelligent User Interfaces*, pp. 60–67. ACM, 2000. doi: 10.1145/325737.325782

[13] A. Dix, J. Finlay, G. Abowd, and R. Beale. *Human-Computer Interaction*. Pearson Education, 3rd ed., 2004.

[14] R. Engels. Planning tasks for knowledge discovery in databases performing task-oriented user-guidance. In *Proc. of Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 170–175. AAAI Press, 1996.

[15] I. Fujishiro, Y. Takeshima, Y. Ichikawa, and K. Nakamura. GADGET: Goal-oriented application design guidance for modular visualization environments. In *Proc. of IEEE Conference on Visualization*, pp. 245–252. IEEE, 1997. doi: 10.1109/VISUAL.1997.663889

[16] S. Gladisch, H. Schumann, and C. Tominski. Navigation recommendations for exploring hierarchical graphs. In *Proc. of Symposium on Visual Computing*, pp. 36–47. Springer, 2013. doi: 10.1007/978-3-642-41939-3_4

[17] D. Gotz and Z. Wen. Behavior-driven visualization recommendation. In *Proc. of Conference on Intelligent User Interfaces*, pp. 315–324. ACM, 2009. doi: 10.1145/1502650.1502695

[18] S. Gratzl, N. Gehlenborg, A. Lex, H. Pfister, and M. Streit. Domino: Extracting, comparing, and manipulating subsets across multiple tabular datasets. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2023–2032, 2014. doi: 10.1109/TVCG.2014.2346260

[19] C. Healey and B. Dennis. Interest driven navigation in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(10):1744–1756, 2012. doi: 10.1109/TVCG.2012.23

[20] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. *Information Visualization (SAGE)*, 7(1):49–62, 2008. doi: 10.1057/palgrave.ivs.9500167

[21] Y. Hijikata. Implicit user profiling for on demand relevance feedback. In *Proc. of Conference on Intelligent User Interfaces*, pp. 198–205. ACM, 2004. doi: 10.1145/964442.964480

[22] E. Horvitz. Principles of mixed-initiative user interfaces. In *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, pp. 159–166. ACM, 1999. doi: 10.1145/302979.303030

[23] T. Jankun-Kelly and K.-L. Ma. A spreadsheet interface for visualization exploration. In *Proc. of IEEE Conference on Visualization*, pp. 69–76. IEEE, 2000. doi: 10.1109/VISUAL.2000.885678

[24] D. Koop, C. Scheidegger, S. Callahan, J. Freire, and C. Silva. VisComplete: Automating suggestions for visualization pipelines. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1691–1698, 2008. doi: 10.1109/TVCG.2008.174

[25] J. Krause, A. Perer, and E. Bertini. Infuse: interactive feature selection for predictive modeling of high dimensional data. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1614–1623, 2014. doi: 10.1109/TVCG.2014.2346482

[26] M. Kreuseler, T. Nocke, and H. Schumann. A history mechanism for visual data mining. In *Proc. of IEEE Symposium on Information Visualization*, pp. 49–56. IEEE, 2004. doi: 10.1109/INFVIS.2004.2

[27] D. J. Lehmann, F. Kemmler, T. Zhyhalava, M. Kirschke, and H. Theisel. Visualnostics: Visual guidance pictograms for analyzing projections of high-dimensional data. *Computer Graphics Forum*, 34(3):291–300, 2015. doi: 10.1111/cgf.12641

[28] A. Lex, M. Streit, H.-J. Schulz, C. Partl, D. Schmalstieg, P. J. Park, and N. Gehlenborg. StratomeX: Visual analysis of large-scale heterogeneous genomics data for cancer subtype characterization. *Computer Graphics Forum*, 31(3pt3):1175–1184, 2012. doi: 10.1111/j.1467-8659.2012.03110.x

[29] M. Luboschik, C. Maus, H.-J. Schulz, H. Schumann, and A. Uhrmacher. Heterogeneity-based guidance for exploring multiscale data in systems biology. In *Proc of IEEE Symposium on Biological Data Visualization*, pp. 33–40. IEEE, 2012. doi: 10.1109/BioVis.2012.6378590

[30] T. May, M. Steiger, J. Davey, and J. Kohlhammer. Using signposts for navigation in large graphs. *Computer Graphics Forum*, 31(3pt2):985–994, 2012. doi: 10.1111/j.1467-8659.2012.03091.x

[31] M. Maybury. Discovering distributed expertise. In *Proc. of AAAI Fall Symposium on Regarding the Intelligence in Distributed Intelligent Systems*, pp. 100–103. AAAI, 2007.

[32] M. Maybury, R. D'Amore, and D. House. Expert finding for collaborative virtual environments. *Communications of the ACM*, 44(12):55–56, 2001. doi: 10.1145/501317.501343

[33] D. Norman. *The Design of Everyday Things*. Basic Books, 2013.

[34] D. Oard and J. Kim. Implicit feedback for recommender systems. In *Proc. of AAAI workshop on Recommender Systems*, pp. 81–83. AAAI, 1998.

[35] http://www.oxforddictionaries.com/definition/english/guidance. Accessed: March 31st, 2016.

[36] http://www.merriam-webster.com/dictionary/guidance. Accessed: March 31st, 2016.

[37] A. Perer and B. Shneiderman. Systematic yet flexible discovery: guiding domain experts through exploratory data analysis. In *Proc. of Conference on Intelligent User Interfaces*, pp. 109–118. ACM, 2008. doi: 10.1145/1378773.1378788

[38] H. C. Purchase, E. Hoggan, and C. Görg. How important is the "mental map"? – an empirical investigation of a dynamic graph layout algorithm. In *Graph Drawing*, pp. 184–195. Springer, 2007. doi: 10.1007/978-3-540-70904-6_19

[39] A. Radloff, G. Fuchs, and H. Schumann. Supporting visual analysis in smart meeting rooms. In *Proc. of Workshop on Visual Analytics*, pp. 1–5. Eurographics, 2012. doi: 10.2312/PE/EuroVAST/EuroVA12/001-005

[40] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. A. Keim. Knowledge generation model for visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1604 – 1613, Dec. 2014. doi: 10.1109/TVCG.2014.2346481

[41] H.-J. Schulz, T. Nocke, M. Heitzler, and H. Schumann. A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2366–2375, 2013. doi: 10.1109/TVCG.2013.120

[42] H.-J. Schulz, M. Streit, T. May, and C. Tominski. Towards a characterization of guidance in visualization. Poster at IEEE Conference on Information Visualization (InfoVis), 2013.

[43] Y. Shrinivasan and J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proc. of SIGCHI Conference on Human Factors in Computing Systems*, pp. 1237–1246. ACM, 2008. doi: 10.1145/1357054.1357247

[44] S. Smith and J. Mosier. Guidelines for designing user interface software. Technical Report ESD-TR-86-278, Mitre Corporation, Bedford MA, 1986.

[45] M. Streit, H.-J. Schulz, A. Lex, D. Schmalstieg, and H. Schumann. Model-driven design for the visual analysis of heterogeneous data. *IEEE Transactions on Visualization and Computer Graphics*, 18(6):998–1010, 2012. doi: 10.1109/TVCG.2011.108

[46] J. Thomas and K. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, 2005.

[47] J. van Wijk. Views on visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(4):421–433, 2006. doi: 10.1109/TVCG.2006.80

[48] Y. Wang, Q. Shen, D. Archambault, Z. Zhou, M. Zhu, S. Yang, and H. Qu. AmbiguityVis: Visualization of ambiguity in graph layouts. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):359–368, 2016. doi: 10.1109/TVCG.2015.2467691

[49] P. Wegner. Why interaction is more powerful than algorithms. *Communications of the ACM*, 40(5):80–91, 1997. doi: 10.1145/253769.253801

[50] S. Wehrend and C. Lewis. A problem-oriented classification of visualization techniques. In *Proc. of IEEE Conference on Visualization*, pp. 139–143. IEEE, 1990. doi: 10.1109/VISUAL.1990.146375

[51] W. Willett, J. Heer, and M. Agrawala. Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136, 2007. doi: 10.1109/TVCG.2007.70589

[52] M. Wohlfart and H. Hauser. Story telling for presentation in volume visualization. In *Proc. of Eurographics Conference on Visualization*, pp. 91–98. Eurographics, 2007. doi: 10.2312/VisSym/EuroVis07/091-098

[53] K. Wongsuphasawat, D. Moritz, A. Anand, J. Mackinlay, B. Howe, and J. Heer. Voyager: Exploratory analysis via faceted browsing of visualization recommendations. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):649–658, 2016. doi: 10.1109/TVCG.2015.2467191

[54] D. Yang, Z. Xie, E. Rundensteiner, and M. Ward. Managing discoveries in the visual analytics process. *SIGKDD Explor. Newsl.*, 9(2):22–29, 2007. doi: 10.1145/1345448.1345453

[55] L. Yu, A. Lu, W. Ribarsky, and W. Chen. Automatic animation for time-varying data visualization. *Computer Graphics Forum*, 29(7):2271–2280, 2010. doi: 10.1111/j.1467-8659.2010.01816.x