The Role of Explicit Knowledge: A Conceptual Model of Knowledge-Assisted Visual Analytics

Paolo Federico,*† Markus Wagner,*‡ Alexander Rind,‡ Albert Amor-Amorós,† Silvia Miksch,† Wolfgang Aigner‡

ABSTRACT

Visual Analytics (VA) aims to combine the strengths of humans and computers for effective data analysis. In this endeavor, humans' tacit knowledge from prior experience is an important asset that can be leveraged by both human and computer to improve the analytic process. While VA environments are starting to include features to formalize, store, and utilize such knowledge, the mechanisms and degree in which these environments integrate explicit knowledge varies widely. Additionally, this important class of VA environments has never been elaborated on by existing work on VA theory. This paper proposes a conceptual model of Knowledge-assisted VA conceptually grounded on the visualization model by van Wijk. We apply the model to describe various examples of knowledge-assisted VA from the literature and elaborate on three of them in finer detail. Moreover, we illustrate the utilization of the model to compare different design alternatives and to evaluate existing approaches with respect to their use of knowledge. Finally, the model can inspire designers to generate novel VA environments using explicit knowledge effectively.

Keywords: Automated analysis, tacit knowledge, explicit knowledge, visual analytics, information visualization, theory and model.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Theory and methods.

1 INTRODUCTION

Analytical reasoning for real world decision making involves volumes of uncertain, complex, and often conflicting data that analysts need to make sense of. In addition to sophisticated analysis methods, knowledge about the data, the domain, and prior experience are required to not get overwhelmed in this endeavor. Ideally, a Visual Analytics (VA) environment would leverage this knowledge to better support domain users, their data, and the analytical tasks in context.

Let us examine the role of knowledge in data analysis in an illustrative scenario from the medical domain: Alice, a medical expert, analyzes patient data. One possible objective of the analysis is a differential diagnosis: Alice needs to interpret data in order to identify a particular critical condition among different candidate conditions. However, different conditions might be present at the same time and, therefore, Alice has also to analyze co-morbidity. After having identified the condition(s), Alice needs to take action and prescribe the best possible therapeutic strategies. Data analysis is used to support evidence-based decision making. Moreover, Alice might need to adapt existing evidence-based therapies, which represent best on-average choices for large populations, to the specific situation of individual patients. In addition, Alice might want to consult other experts and ask them about their opinion or their previous experience with similar cases. In many cases, patients are also involved in a shared decision. Alice informs patients about the possible options and their consequences. She supports them to make better informed decisions while taking into account individual preferences. Afterwards, Alice might perform follow-up or retrospective analysis in order to check the compliance to the therapeutic plans as well as their effectiveness; the objective is the iterative refinement of evidence-based diagnosis and therapy.

All the phases of this example scenario involve prior knowledge. Alice relies on her prior knowledge to select appropriate analytical methods and to interpret the results. For decision-making, she exploits her knowledge of evidence-based therapy, knowledge about similar cases, and knowledge from other experts. Moreover, Alice has to fill knowledge gaps with her patient in shared-decision making.

Supporting such complex scenarios by explicitly taking advantage of expert knowledge in a VA system gives rise to more effective environments for gaining insights. I.e., making use of auxiliary information about data and domain specifics in addition to the raw data, will help to better select, tailor, and adjust appropriate methods for visual representation, interaction, and automated analysis.

To facilitate such epistemic processes, a number of visualization researchers have repeatedly called for the integration of knowledge with visualization. Chen [18] argues that visualization systems need to be adaptive for accumulated knowledge of users, especially domain knowledge needed to interpret results. A specific recommendation in the research and development agenda for VA by Thomas and Cook prescribes to "develop knowledge representations to capture, store, and reuse the knowledge generated throughout the entire analytic process" [72, p.42]. In their discussion of the science of interaction, Pike et al. [57] point out that VA environments have only underdeveloped abilities to represent and reason with human knowledge. Therefore, they declare knowledge-based interfaces as one of seven research challenges. Even a special issue of the journal IEEE Computer Graphics and Applications was dedicated to knowledge-assisted visualization [21]. These calls have resulted in a number of visualization environments that include features to generate, transform, and utilize explicit knowledge. However, the mechanisms and degree to which these environments integrate explicit knowledge vary widely. Additionally, this important class of VA environments has not yet been investigated from a more systematic, conceptual perspective of VA theory. This raises the need for a knowledge-assisted VA model describing the integration of explicit knowledge, its extraction and its application in the VA process. Such a model could act as means for systematically discussing knowledge-assisted VA approaches, comparing and relating them, as well as being used as system blueprint to design novel VA systems.

In this paper, we aim to fill this gap in theory by systematically investigating the role of explicit knowledge in VA, by proposing a model for knowledge-assisted VA, and by demonstrating its application. The main contributions of our work are to:

- provide a conceptual abstraction and theoretical modeling of VA processes based on the introduction of our novel knowledge-assisted VA model (Section 3).
- illustrate the possibilities of explicit knowledge integration and

*Paolo Federico and Markus Wagner equally contributed to this paper and are both to be regarded as first authors.

†Paolo Federico, Albert Amor-Amorós, and Silvia Miksch are with TU Wien, Austria. E-mail: {federico, amor, miksch}@ifs.tuwien.ac.at

‡Markus Wagner, Alexander Rind, and Wolfgang Aigner are with St. Poelten University of Applied Sciences, Austria and TU Wien, Austria. E-mail: {markus.wagner, alexander.rind, wolfgang.aigner}@fhstp.ac.at

extraction, the integration of automated data analysis methods as well as the combination of both (Section 3).
- demonstrate the utility of the model in Section 4 as its ability: 1) to describe the functionalities of existing approaches and to categorize them in relation of the included components and processes; 2) to express the costs and benefits of knowledge-assisted processes and systems; and 3) to inspire new research directions and to enable design of innovative approaches.

2 Background and Related Work

In this section we are presenting a general view of the role of knowledge in visualization (see Section 2.1), followed by a detailed presentation of well-known models describing visualization in several levels of detail (see Section 2.2), and how knowledge is integrated and supported.

2.1 Knowledge in Visualization

Discovery, acquisition, and generation of new knowledge are main aims of VA. According to Thomas and Cook [72, p. 42] the final task of the analytical reasoning process is to create some kind of knowledge product or direct action based on gained insight. Both, interactive visualization and automated data analysis, whose combination has been defined VA [45], share the same aim. Information visualization aims at amplifying human cognition [15] or, in other words, mental action or process of acquiring knowledge and understanding; analogously, the aim of automated data analysis methods is, by definition, knowledge discovery [30].

The meaning of terms such as data, information, and knowledge, as well as the ways they relate to each other, are widely but often inconsistently used. In the field of visualization, Chen at al. [19] untangle the terminology, deriving it from the data-information-knowledge-wisdom (DIKW) pyramid. While the inspiration of the DIKW pyramid has been traced back to verses by T.S.Eliot [60], slightly different versions have been proposed in different domains, for example in information sciences by Ackoff [1] and in knowledge management by Zeleny [80]; different versions sometimes include three items only (earlier variants omit data, later variants omit wisdom), and some introduce additional items (e.g., understanding between knowledge and wisdom, or enlightenment beyond wisdom). However, an aspect that all the different formulations have in common is that the levels of structure, meaning, value, and/or human agency increase from data to wisdom [60]. Chen et al. [19] do not focus on structural differences but on the functional differences outlined by Ackoff [1] and omit wisdom; they describe data as symbols, information as data that are processed to be useful, providing answers to “who”, “what”, “where”, and “when” questions, and knowledge as application of data and information, providing answers to “how” questions. Other authors describe knowledge in the context of the DIKW pyramid as a combination of data and information, complemented with expert opinion, skills, experience, expertise, and accumulated learning. This can be applied to a particular problem or activity and can be used to aid decision making and predispose people to act in a particular way [60]. Moreover, Chen et al. [19] also observe that data, information, and knowledge are processed by both humans and computers and, therefore, they extend their meanings from the cognitive and perceptual space to the computational space; in particular, they define knowledge in the computational space as “data that represents the results of a computer-simulated cognitive process, such as perception, learning, association, and reasoning, or the transcripts of some knowledge acquired by human beings” [19, p. 13].

The distinction between the cognitive and perceptual (i.e., human) space, on the one hand, and the computational (i.e., machine) space, on the other hand, was also applied by Wang et al. [78]. They distinguish between **tacit** knowledge and **explicit** knowledge: tacit knowledge can be understood as knowledge which users hold in their minds, it is personal and specialized, and it can only be acquired by humans through their cognitive processes; explicit knowledge has been written, saved, or communicated and, therefore, can be stored in a database and processed by a computer.

In the human cognition process, new knowledge is gained by establishing relations between new insights and prior knowledge, deriving from previous experience or learning. In particular, two types of **prior** knowledge are needed by a user to understand the intended message in visualization: **operational** knowledge (how to interact with the information visualization system), and **domain** knowledge (how to interpret the content) [18]. While a focus on usability and a perception- and cognition-aware design can alleviate the need for operational knowledge, the domain knowledge cannot be easily replaced [18]. Thus, the research on the problem of operational knowledge in visualization has focused on the science of interaction: Pike et al. [57] identify the design of *knowledge-based interfaces* as an open challenge, stating that the ability of visual analysis tools to represent and reason with human knowledge is underdeveloped. Knowledge-assisted visualization aims at exploiting both types of knowledge: sharing domain knowledge among different users and reducing the operational knowledge needed by users of complex visualization techniques [21].

According to Thomas and Cook [72], the proper representation of final as well as intermediate **generated** knowledge can be useful to support the analytical discourse, the interoperation between its human and machine components, and the collaboration between different users, as well as to trace the relations between data and derived knowledge products, by retaining quality and provenance information.

Automated analysis methods can also benefit greatly from the use of prior knowledge. In fact, the fundamental role of prior knowledge in the knowledge discovery process (KDD) has been already emphasized more than 20 years ago [30]. Intelligent data analysis, or the application of artificial intelligence (AI) techniques in data analysis, aims at automatically extracting information from data by exploiting explicit domain knowledge (sometimes called background knowledge in this context) [40]. Knowledge-based systems enable the integration of explicit knowledge into the reasoning process, so that it is easy to model exceptional rules, which for example can prevent the system to reason over abnormal conditions [56]. Novel approaches for knowledge-based data analysis and interpretation using computer-readable explicit knowledge have obvious advantages over those that do not [81]. Prior knowledge, for example, can be used to specify appropriate features or techniques, or provide a representation of the output that is easy to interpret.

In summary, by assessing the role of knowledge in visualization, besides untangling concepts and terminology, we observe several calls to investigate ways to integrate both prior knowledge and intermediate knowledge products in the VA discourse, by adequate representation and processing as well as diverse approaches towards this direction.

2.2 Models in Visualization

Even though knowledge plays such a central role, existing models of visual data analysis involve the notion of knowledge to varying extents.

The classical *visualization pipeline* [14, 15] as well as the *data flow model* and the *data state model* [22–24] do not mention knowledge explicitly. Still we can assume that they imply it: first, visualization is aimed to amplify cognition, i.e., the mental process of knowledge acquisition; second, interactive transformations at any stage of the pipeline allow intervention of users’ domain knowledge and require their operational knowledge.

Van Wijk [74] propose an operational model of visualization in order to describe the context in which visualization operates and characterize its value. The model identifies three spaces: the data space, the visualization space (i.e., the machine space), and the user space. Moreover, the model explicitly includes knowledge within the user space and involves it in two dynamic processes: existing knowledge is involved in the perception/cognition process in order to gain new knowledge about data from the visualization, as well as in the exploration process to specify the visualization algorithms and parameters. Van Wijk's model has been broadly adopted, critiqued, and extended by visualization scholars. Green et al. [37] propose a human cognition model for VA and relate it to the simple model of visualization by van Wijk, by observing that perception, knowledge, and exploration should be all modeled as cognitive processes informing each other. Wang et al. [78] extend van Wijk's model by adding a knowledge base that contains explicit knowledge and uses it to describe four knowledge conversion processes: *internalization*, by which a user continuously builds tacit (internal) knowledge based on perceptually, cognitively, and interactively incorporating the visualized explicit knowledge; *externalization*, by which internally created tacit knowledge can be extracted and saved into the knowledge base; *collaboration*, by which distinct users share tacit knowledge by using visualization or by direct communication; and *combination*, by which new explicit knowledge can be combined with existing explicit knowledge in a knowledge base. Ceneda et al. [17] build upon van Wijk's model to characterize guidance in VA. They consider explicit domain knowledge and user knowledge as inputs to the guidance process, together with the data and the full specification history; however, while the domain knowledge is explicit, they do not detail the processes by which a user's tacit knowledge can be externalized and made available for guidance on the computer side.

The *sensemaking loop* by Pirolli and Card [58] is based on a hierarchy of representations with increasing levels of structure and human effort: information, schema, insight, and product. Besides a different terminology, this hierarchy is similar to the DIKW pyramid and its final outcome is a knowledge product as recommended by Thomas and Cook [72]. However, this model does not describe the analytical discourse between the machine and the human as well as the cognitive processes of the latter in detail; neither are the role of prior knowledge and in particular explicit knowledge considered.

The process of knowledge discovery in databases (KDD) as modeled by Fayyad [30] consists of subsequent steps (selection, preprocessing, transformation, and data mining) which produce increasingly elaborated artifacts from raw data up to patterns which, at the final step, need to be evaluated and interpreted by the user in order to gain new knowledge. A limitation of the model, also recognized by its authors, is the lack of adequate means to integrate and utilize prior knowledge in the process. The *visual KDD* model [39] addresses this problem by combining a KDD pipeline with an interactive visualization pipeline, but the processes involving knowledge, both on the human side and on the computer side, are not detailed.

The model of the *VA process* by Keim et al. [45,46] combines automated analysis methods with human interaction to gain knowledge or insights from the data. In this model, intermediate knowledge products are denoted as hypotheses and analytical models, and the only considered knowledge is the knowledge that the user acquires by perception and cognition; moreover, explicit knowledge is not included and the only way to integrate prior knowledge is by interaction loops. Lammarsch et al. [50] developed an extension of the VA process model, including explicitly domain knowledge about time-oriented data obtained from previous analyses [46]. Another extension is the *knowledge generation model* by Sacha et al. [62]. It elaborates human interaction with the VA environment as three loops producing increasingly meaningful artifacts called finding, insight, and knowledge. However, these artifacts are situated solely on the human side.

Ribarsky and Fisher [59] extend the knowledge generation model by Sacha et al. [62] even further, proposing a *human-machine interaction loop* similar to the sensemaking loop by Pirolli and Card [58]. In particular, on the computer side, they add prior knowledge, i.e., explicit knowledge derived from external knowledge, from previous analysis sessions of the same user, or from collaboration with other users; on the human side, they add user knowledge, i.e., knowledge from education and past experience that the user carries into the knowledge generation, synthesis, and hypothesis-shaping processes. It is worth noting that Ribarsky and Fisher explicitly denote the analytical models as pieces of explicit knowledge, while hypotheses are placed between tacit and explicit knowledge.

The models described above demonstrate the different properties knowledge can present and the different roles knowledge can play in the VA context. While each model emphasizes interesting aspects, none of them covers them all.

3 Conceptual Model of Knowledge-Assisted VA

As discussed in the previous section, knowledge in the VA process is not sufficiently addressed by the existing models. To fill this gap, we propose a model for knowledge-assisted VA. First, we describe the requirements that such a model needs to meet. Following that, the model is constructed and formally described. Finally, the involved knowledge dimensions are discussed.

3.1 Eliciting Model Criteria

By deriving general characteristics from the analysis of single models in Section 2.2, we claim that a unified model of knowledge-assisted VA should be able to capture different VA components, spaces, knowledge types, and knowledge processes.

VA can be understood as a combination of automatic analysis, visualization, and interaction methods and, therefore, these three **VA components** need to be modeled. Models developed for visualization can lack an analysis component, while models developed for KDD do not take visualization into sufficient consideration. However, also some models developed for VA can disregard the representation of these components. The model for knowledge-assisted VA by Wang et al. [77], for example, inherits a visualization and an interactive specification component by van Wijk [74], but does not expressly include an analysis component; the model of sensemaking by Pirolli and Card [58] does not distinguish among these components at all.

As for **spaces**, many models distinguish between the conceptual and perceptual space (on the human side) and a computational space (on the machine side). This articulation is useful, since it allows us to describe and exploit perception and cognition processes but also to design and validate system features and algorithms. Moreover, it enables a representation of the analytical discourse as a collaboration between user and computer, including important processes across the human-machine interface.

A good model for knowledge-assisted visualization has to include different **knowledge types**, namely domain and operational knowledge as well as tacit and, most important, explicit knowledge (obtained either by externalization of the user's tacit knowledge, or by a computer-simulated cognitive process). Tacit knowledge is obviously involved in all human cognition processes, while the integration of explicit knowledge is the added value of knowledge-assisted approaches.

An operational model should capture the mechanisms behind the different **knowledge processes** as dynamic phenomena, whose current state depends on an initial state and on the full history. This representation better reflects the epistemic nature of knowledge acquisition, which is an accumulative phenomenon – new knowledge is generated by relating new insight with prior knowledge.

While none of the afore-mentioned models fulfills all these criteria, we can obtain a general model by extending an existing model. A good candidate is the simple model of visualization by van Wijk [74]. It clearly distinguishes between the human and the computer space,

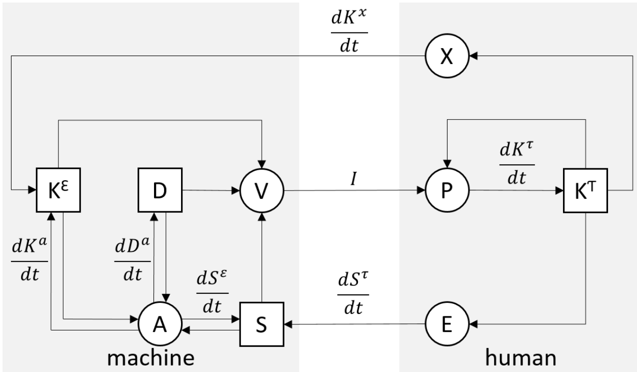

Figure 1: Conceptual Model of Knowledge-Assisted VA. The model is divided into two spaces (machine and human) and describes knowledge generation, conversion, and exploitation within the VA discourse, in terms of processes: analysis A, visualization V, externalization X, perception/cognition P, and exploration E; containers: explicit knowledge $\boxed{K^\varepsilon}$, data $\boxed{D}$, specification $\boxed{S}$, and tacit knowledge $\boxed{K^\tau}$; and a non-persistent artifact: image I.

and it is an operational model which effectively describes knowledge processes and loops, on the human side only. Its original version, indeed, does not include explicit knowledge, but Wang et al. [77] have shown that it can be extended in this sense. However, both the original and the extended version do not expressly represent the different components of VA and need to be properly adapted.

3.2 Constructing the Model

For developing our model (see Figure 1), we use the formalism introduced by van Wijk [74] to describe the operational context of visualization: circles represent processes, and boxes represent containers where input and outputs are continuously accumulated and accessed. In particular, van Wijk's model encompasses the following processes: on the machine space, visualization V; on the human space, perception and cognition P, and exploration E. The following containers are also involved, since they constitute inputs or outputs to one or more processes: on the machine space, data $\boxed{D}$, and specification $\boxed{S}$; on the human space, tacit (see Section 2.1) knowledge $\boxed{K^\tau}$. In order to capture the role of explicit knowledge in VA, we need to incorporate two additional elements, which lie on the machine space: on the one hand, a container accounting for the existence of explicit knowledge itself, $\boxed{K^\varepsilon}$; on the other hand, a process that accounts for the existence of automatic analysis methods A, a defining characteristic of VA approaches.

In the following, we elaborate on the construction of our model by eliciting three different types of processes involving the use of tacit and explicit knowledge, namely knowledge generation, conversion, and exploitation. The reader must be aware of the fact that, even though we discuss these processes individually to justify the construction of our model in a systematic way, they generally occur together and, more importantly, their effectiveness depends on their combined action. In addition to the graphical and textual description provided below, the supplemental material includes a formal mathematical description of the model's processes and containers.

3.2.1 Knowledge Generation

We begin the construction of the model by considering the processes involved in the generation of knowledge from data.

In van Wijk's model, visualization V is defined as the transformation of data $\boxed{D}$ into an image I given a certain specification $\boxed{S}$; tacit knowledge $\boxed{K^\tau}$ is generated on the basis of that image through humans' perceptual and cognitive abilities P.

$$\begin{array}{ccccccc} \boxed{D} & \longrightarrow & V & \xrightarrow{I} & P & \longrightarrow & \boxed{K^\tau} \\ & & \uparrow & & & & \\ & & \boxed{S} & & & & \end{array}$$

In a VA scenario [8, 46], the visualization pipeline is complemented with an **automated data analysis** pipeline, supporting knowledge generation with automatic methods A aimed at the elicitation of explicit knowledge $\boxed{K^\varepsilon}$ in its different forms (e.g., models, rules, parameter settings), given a certain specification $\boxed{S}$.

$$\begin{array}{ccccc} \boxed{D} & \rightarrow & A & \rightarrow & \boxed{K^\varepsilon} \\ & & \uparrow & & \\ & & \boxed{S} & & \end{array}$$

3.2.2 Knowledge Conversion

Our model should also encompass the transformation of explicit knowledge into tacit knowledge (i.e., knowledge transfer from the machine to the human), as well as that of tacit knowledge into explicit knowledge (i.e., knowledge transfer from the human to the machine). Wang [78] refers to the former as knowledge internalization and to the latter as knowledge externalization.

Knowledge **internalization** is required when explicit knowledge is automatically extracted from data, or when an external source of explicit knowledge is being used. In some cases, it is performed directly, i.e., through **knowledge visualization** [51, 71] in terms of the concepts and relationships involved, by considering them a specific form of data:

$$\boxed{K^\varepsilon} \longrightarrow V \xrightarrow{I} P \longrightarrow \boxed{K^\tau}$$

In other cases, knowledge internalization occurs indirectly, i.e., through the generation and subsequent visualization of datasets that provide users with the scenarios that result from the application of the knowledge (e.g., **simulation**):

$$\boxed{K^\varepsilon} \longrightarrow A \longrightarrow \boxed{D} \longrightarrow V \xrightarrow{I} P \longrightarrow \boxed{K^\tau}$$

Knowledge **externalization**, on the other side, is required when an explicit representation of the user's tacit knowledge is needed. In some scenarios, the system might support the user in actively formulating that knowledge through an appropriate **direct externalization** interface X:

$$\boxed{K^\tau} \longrightarrow X \longrightarrow \boxed{K^\varepsilon}$$

In other cases, tacit knowledge can be automatically inferred from the user's sensemaking process and domain expertise by methods for **interaction mining** (e.g., semantic interaction analysis [28]):

$$\boxed{K^\tau} \longrightarrow E \longrightarrow \boxed{S} \longrightarrow A \longrightarrow \boxed{K^\varepsilon}$$

3.2.3 Knowledge Exploitation

Knowledge is, for obvious reasons, generally regarded as the ultimate outcome of the analytical process. However, knowledge generation typically relies on knowledge exploitation to boost its effectiveness. In other words, knowledge also plays a fundamental role as an input to any VA workflow.

On the human space, the feedback mechanisms by which tacit knowledge supports both perception and cognition P as well as interactive exploration E are captured in van Wijk's model:

$$\begin{tikzcd} \boxed{D} \arrow[r] & V \arrow[r, "I"] & P \arrow[r] & \boxed{K^\tau} \arrow[l, bend right] \arrow[dl, bend left] \\ & \boxed{S} \arrow[u] & E \arrow[l] & \end{tikzcd}$$

Analogous exploitation mechanisms for explicit knowledge appear on the machine space: the fundamental importance of prior knowledge to the KDD process has already been recognized [30], and the term **intelligent data analysis** [40] has been coined for referring to the use of explicit knowledge in order to improve existing automatic knowledge extraction methods.

$$\begin{tikzcd} \boxed{D} \arrow[r] & A \arrow[r] & \boxed{K^\varepsilon} \arrow[l, bend right] \\ & \boxed{S} \arrow[u] & \end{tikzcd}$$

Moreover, explicit knowledge can also be leveraged to provide **guidance** [17]. Inputs for guidance are explicit knowledge $\boxed{K^\varepsilon}$, data $\boxed{D}$, and specification $\boxed{S}$ (containing the full history of previous settings interactively explored by the user to specify images), which are analyzed A to generate specification suggestions. These suggestions can be used automatically, or combined with user interactive exploration E in the context of mixed-initiative systems [41].

$$\begin{tikzcd} & V \arrow[r, "I"] & P \arrow[r] & \boxed{K^\tau} \arrow[dl, bend left] \\ \boxed{D} \arrow[ur, bend left] \arrow[dr, bend right] & \boxed{S} \arrow[u] \arrow[d, shift left] & E \arrow[l] & \\ & A \arrow[u, shift left] \arrow[r] & \boxed{K^\varepsilon} \arrow[l, bend left] & \end{tikzcd}$$

3.3 Characterizing the Analysis processes

The formalism we adopted is general enough to model knowledge-assisted VA at a high level of abstraction. For a finer-grained modeling, both processes and containers can be broken down into sub-components and characterized in detail. In particular, the automated analysis process A can be understood as an aggregation of different algorithmic methods, namely guidance G, simulation U, and data mining/machine learning M (see Figure 2). The guidance process G encompasses different techniques, that have been classified according to domain, input, output, type, and degree [17]. The simulation process U comprehends diverse algorithmic methods that can be used to synthesize new data starting from explicit knowledge. The data mining/machine learning M is directly involved in the knowledge generation process and can support common KDD tasks [30]: classification, regression, clustering, summarization, association rule learning, and anomaly detection. Instead of the raw data $\boxed{D}$, we can use the entire specification data store $\boxed{S}$ as an input to M: this is the case of the interaction mining process, with its specific methods (e.g., semantic interaction [28]). However, the detailed discussion of all algorithms that are comprised within the analysis process A goes beyond the scope of this paper. Moreover, the model can be easily extended by instantiating specific sub-processes in order to cover possible emerging directions in knowledge-assisted VA.

3.4 Characterizing the Knowledge

Knowledge involved in knowledge-assisted VA can be classified according to several dimensions. In Section 2.1 we have already introduced the distinction between tacit knowledge and explicit

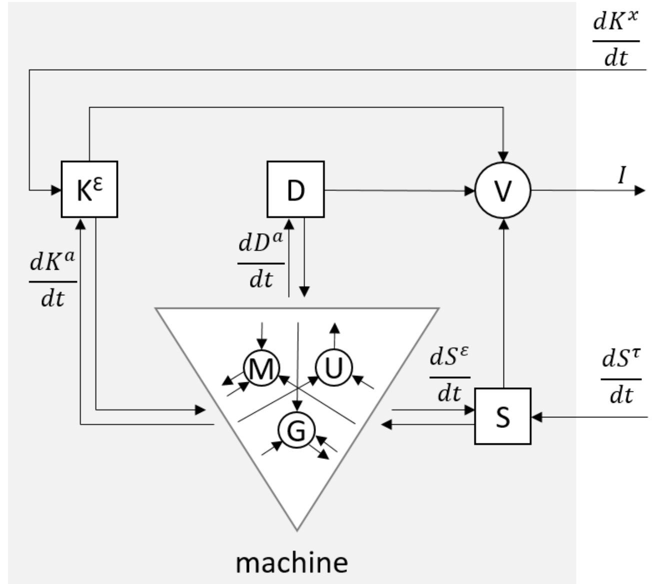

Figure 2: Our conceptual model describes Knowledge-assisted VA at a high level of abstraction; nevertheless, processes can be decomposed into sub-processes, enabling a finer-grained specification. In this close-up figure, analysis A is broken down into three possible components, namely data mining/machine learning M, simulation U, and guidance G.

knowledge. This distinction primarily refers to the **space**: tacit knowledge resides in the cognitive/conceptual space, while explicit knowledge in the computational space. From the human perspective, tacit knowledge is internal knowledge, while explicit knowledge has been externalized. However, there might be cases of externalized knowledge that is not directly computer-interpretable, for example annotations by natural language or free drawings, requiring a preliminary mining to be made available to further computational steps as explicit knowledge.

In Section 2.1 we have also discussed the **type**: knowledge is either operational knowledge or domain knowledge. An additional dimension is the representation paradigm: following the classical distinction used in AI, we distinguish between declarative and procedural knowledge. In short, declarative (also, descriptive) knowledge is the knowledge of *what*, while procedural (or imperative) knowledge is the knowledge of *how* and *how best*. The former has a focus on data and information, the latter on procedures. Both declarative and procedural knowledge belong into domain knowledge; in principle, the former can help users make sense of data, the latter make decisions and take action in the application domain. Nevertheless, procedural knowledge can also be used for retrospective analysis (e.g., checking if the undertaken decisions were correct). It is worth noting, however, that procedural knowledge is domain knowledge, supporting domain-specific reasoning, and must not be confusing with operational knowledge, which the user needs to operate the VA environment.

Furthermore, knowledge can be classified according to its **origin**, comprising the source it comes from and the time it is made available, with respect to the design and the use of the VA environment. Knowledge can exist prior to the VA environment, e.g., if it has been collected and formalized in the application domain independently of the environment at hand. Knowledge can be acquired and specified on purpose when a VA environment is designed and implemented, by designers and knowledge engineers. Finally, knowledge can be generated during the environment's operation, either from data, or

Table 1: Examples of knowledge-assisted visual analytics classified after our model

			Finding Waldo [13]	Knave/Visitors [48]	Smart Grids [67]	KEGS [79]	IMAGE [53]	Compliance [9]	EVE [11]	SemViz [36]	Compliance [5]	Kav-db [34]	Dabek et al. [27]	VisExemplar [63]	Prajna [69]	Bio ontology [16]	Qualizon Graphs [32]	SemTimeZoom [3,4]	Garg et al. [35]	Smart superviews [52]	DEL [12]	CareCruiser [38]	CareVis [2]	Nam et al. [55]	PORGY [73]	RuleBeneder [66]	VisPad [65]	Sport Events [25]	KAVAGait	VizAssist [10]	Kamsu et al. [44]	Gnaeus [33]	KAMAS [75]	FMVAS [54]
Process	Data Analysis:	$\boxed{D} \rightarrow A \rightarrow \boxed{K^\varepsilon}$							●	●														●										
	Knowledge visualization:	$\boxed{K^\varepsilon} \rightarrow V$									●					●	●	●			●	●	●	●	●	●	●	●	●			●	●	●
	Simulation:	$\boxed{K^\varepsilon} \rightarrow A \rightarrow \boxed{D}$					●																		●	●						●		
	Direct externalization:	$\boxed{K^\tau} \rightarrow X \rightarrow \boxed{K^\varepsilon}$							●			●							●	●					●	●	●	●	●	●			●	
	Interaction mining:	$\boxed{S} \rightarrow A \rightarrow \boxed{K^\varepsilon}$	●									●	●	●													●			●	●			
	Intelligent data analysis:	$(\boxed{D}, \boxed{K^\varepsilon}) \rightarrow A \rightarrow \boxed{K^\varepsilon}$		●	●			●			●				●				●		●							●	●		●	●	●	
	Guidance:	$\boxed{K^\varepsilon} \rightarrow A \rightarrow \boxed{S}$				●				●			●	●	●	●	●	●		●		●	●	●						●	●	●	●	●
Type	Operational		●						●	●		●	●	●	●					●				●			●	●		●	●			●
	Domain, declarative			●		●	●	●								●	●	●	●		●	●	●						●			●	●	●
	Domain, procedural				●		●	●			●										●	●	●		●	●						●		●
Origin	Pre-design			●	●		●	●								●	●	●			●	●	●						●			●	●	●
	Design					●				●	●				●	●	●	●		●										●				●
	Post-design, data																		●					●				●			●			
	Post-design, single user		●						●					●					●	●					●	●	●	●	●	●	●		●	●
	Post-design, multiple users											●	●																					

by users. In the latter case, we distinguish between single-user and collaborative multi-user scenarios. Indeed, once explicit knowledge is made available to the VA process, it can be shared in different collaboration scenarios (co-located or distributed, synchronous or asynchronous [43]) as well as for self-collaboration [59].

4 Application of the Model

In the following, we demonstrate that our model can be a useful to the VA community as a theoretical tool. For this, we base our discussion on different goals of visualization theory [7] respectively interaction models [6]: the ability (1) to describe a wide range of existing knowledge-assisted VA approaches, (2) to allow the assessment of design alternatives in terms of costs and profits, and (3) to inspire the design of new approaches and research directions.

4.1 Describing Existing Approaches

First, we illustrate how our model can be used to describe systems from the literature by identifying and naming key concepts. Therefore, we discuss in detail three selected knowledge-assisted VA approaches through the lens of our model and show how this supports a systematic description and comparison thereof.

4.1.1 Survey

We surveyed prototypes and systems in the scientific literature with a focus on, but non limited to, the visualization community. We included all those works where explicit knowledge has a prominent role. The results are summarized in Table 1, which is structured as follows. The 32 surveyed examples are arranged in columns. Rows are broken down into three groups, corresponding to three dimensions of our model: process (introduced in Section 3.2) as well as knowledge type and knowledge origin (introduced in Section 3.4).

Because of our inclusion criterion, all systems include interactive visualization as well as perception/cognition and exploration processes involving tacit knowledge. Therefore, for the sake of simplicity, we have disregarded the space classification distinguishing between tacit and explicit knowledge. For the same reason, we have skipped the common knowledge generation process from the table. After this simplification, the table includes one knowledge generation process (data analysis), knowledge conversion processes (knowledge visualization, simulation, direct externalization, and interaction mining), and knowledge exploitation processes (intelligent data analysis and guidance).

As for type and origin, we observe that operational knowledge is often captured from users by interaction mining [13] and is utilized to generate visual encodings and to provide guidance to users for choosing among them [27, 44]. Users externalize their attributes and preferences by annotation [52], or by assigning scores and rankings [11, 65], also in a multi-user knowledge-sharing scenario [34]. When interaction mining and guidance are tightly integrated, users can also build visualizations by demonstration [63]. However, also pre-existing domain knowledge, in particular declarative knowledge, can be used to guide or automate the choice of visual encodings, by ontology mappings mechanisms [16] and ontology reasoners [36]. Declarative domain knowledge can be also used to analyze data and compute qualitative abstractions for an easier interpretation [3, 32, 48, 79]. Domain knowledge, both declarative and procedural, can be also represented visually [2, 38]. Procedural domain knowledge is often utilized by rule-based engines to automatically analyze data [5, 67]. Rules can exist in the application domain [9, 67], can be elicited by designers [5], edited by users [66, 73], or learned by example [35].

Table 1 provides a general yet accurate overview of existing knowledge-assisted VA systems and demonstrates that our model can effectively describe a wide spectrum thereof. In the following, we illustrate finer details by discussing three selected examples: Gnaeus, KAMAS, and KAVAGait.

4.1.2 Gnaeus: guideline-based healthcare for cohorts

Gnaeus [33] is a guideline-based knowledge-assisted visualization of electronic health records for cohorts (see Figure 3). Evidence-based clinical practice guidelines are sets of statements and recommendations used to improve health care by providing a trustworthy comparison of treatment options in terms of risks and benefits according to patient's status; they condense the complex domain knowledge underneath clinical practice in narrative form. Gnaeus utilizes their formalization as computer-interpretable guidelines (CIGs).

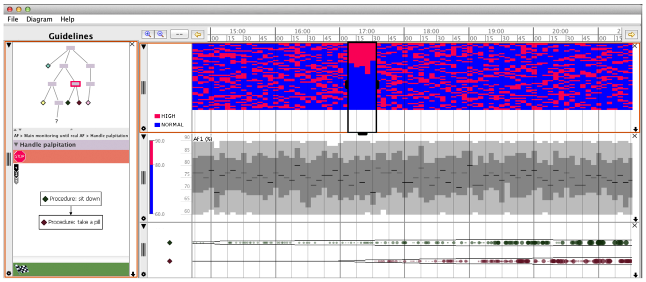

Figure 3: Gnaeus, a guideline-based knowledge-assisted electronic health records visualization for cohorts [33].

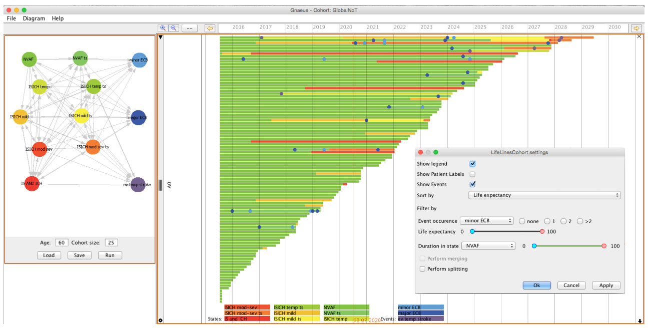

Figure 4: Scipio, a plugin of Gnaeus [33] for simulating patient cohorts.

In Gnaeus, both the **declarative** knowledge and the **procedural** knowledge are exploited to drive two analytical components: the temporal mediator and the compliance analyzer. The declarative knowledge, specified as guideline intentions, is exploited to process the input raw, time-stamped data, such as blood glucose (BG) values at particular times to produce a set of clinically meaningful summarizations and interpretations. The "BG monthly good pattern", for example, is defined as a month when the patient had up to one abnormal value of BG per week and no more than four abnormal values per month, while the BG abnormal values are defined in the context of pregnant diabetic patients according to taking insulin medication and fetus size. Gnaeus computes **knowledge-based temporal-abstraction** (KBTA) [64]: $\{\boxed{D}, \boxed{K^\varepsilon}\} \rightarrow A \rightarrow \boxed{K^\varepsilon}$. To support data interpretation, these qualitative abstractions are visualized together with raw quantitative data by different visual encodings like, for example, qualizon graphs [32]: $\{\boxed{D}, \boxed{K^\varepsilon}\} \rightarrow V$.

Several chronic conditions can be managed with a combination of the right amount of physical activity, appropriate diet, and drugs. Thus, it is particularly important to assess not only the general efficacy of treatments, but also the compliance of patients and caregivers with the clinical guidelines for the management of these diseases. An executed treatment is compliant if the recommendations the patient was eligible for were fulfilled by performing the corresponding actions within the suggested response time windows. In Gnaeus, a rule-based reasoning engine ingests the procedural knowledge of CIGs, patient data, and treatment data, and computes compliance [9]: $\{\boxed{D}, \boxed{K^\varepsilon}\} \rightarrow A \rightarrow \boxed{K^\varepsilon}$, which is then visualized together with raw data $\{\boxed{D}, \boxed{K^\varepsilon}\} \rightarrow V$.

The CIGs are also directly visualized: $\boxed{K^\varepsilon} \rightarrow V$ (**knowledge visualization**). In particular, the hierarchical structure of the guideline is visualized as a tree diagram with a top-down layered layout, whose nodes represent treatment plans and leaves represent clinical actions; the logical structure of a treatment plan is shown as a node-link diagram of a hierarchical task network. Gnaeus also features **knowledge-assisted interactions**, $\boxed{K^\varepsilon} \rightarrow A \rightarrow \boxed{S}$, to support user exploration, $\boxed{K^\tau} \rightarrow E \rightarrow \boxed{S}$.

The Scipio plugin of Gnaeus (see Figure 4) supports shared decision making by interactive visualization of patient-level microsimulation [61]. The evidence-based knowledge about probability of critical event occurrence, as well as transition probabilities between conditions of increasing severity are modeled as Markov models. Since these models might be too complex to be communicated to the patient as such, Scipio utilizes microsimulation to generate data of a synthetic cohort of virtual patients with similar conditions (age, disease, treatment); this data is then visualized for an easier understanding of treatment consequences: $\boxed{K^\varepsilon} \rightarrow A \rightarrow \boxed{D} \rightarrow V$.

4.1.3 KAMAS: behavior-based malware analysis

KAMAS [75] is a knowledge-assisted malware analysis system (see Figure 5). It supports IT-security analysts in learning about previously unknown samples of malicious software (malware) or malware families based on their behavior. Therefore, they need to identify and categorize suspicious patterns from large collections of execution traces. In KAMAS, the analysts are exploring preprocessed call sequences (rules) in their sequential order, containing system and API calls to find out if the observed samples are malicious or not. If a sample is malicious, the system can be used to determine the related malware family. A knowledge database (KDB) **storing explicit knowledge** in the form of rules is integrated into KAMAS to ease the analysis process and to share it with colleagues. Based on the explicit knowledge, **automated data analysis methods** are comparing the rules included in the loaded execution traces based on the specification with the stored explicit knowledge. Thereby, the specification gets adapted to highlight known rules $\{\boxed{D}, \boxed{K^\varepsilon}, \boxed{S}\} \rightarrow A \rightarrow \boxed{S}$. Additionally, the explicit knowledge can be turned on and off partially or completely by interaction: $E \rightarrow \boxed{S}$.

If the analyst loads loaded execution traces into the system, the contained rules are visualized based on the systems specification $\{\boxed{D}, \boxed{S}\} \rightarrow V$. If there is no specification prepared in the first visualization cycle (e.g., zooming, filtering, sorting), all read-in data are visualized and compared to the KDB. The image, which is generated by the visualization process, is perceived by the analyst, **gaining new tacit knowledge** $V \xrightarrow{I} P \rightarrow \boxed{K^\tau}$, which also influences the users perception $\boxed{K^\tau} \rightarrow P$. Depending on the gained tacit knowledge, the analyst has now the ability to interactively explore the visualized malware data (rules) by the system provided methods (e.g., zooming, filtering, sorting), which are affecting the specification $\boxed{K^\tau} \rightarrow E \rightarrow \boxed{S}$. During this interactive process, the analyst gains new tacit knowledge based on the adjusted visualization. For the **integration of new knowledge** into the KDB, the analyst can, on the one hand, add whole rules and on the other hand, the analyst can add a selection of interesting calls, extracting his/her tacit knowledge $\boxed{K^\tau} \rightarrow X \rightarrow \boxed{K^\varepsilon}$. Moreover, KAMAS directly visualizes the whole store explicit knowledge in the KDB $\boxed{K^\varepsilon} \rightarrow V$.

4.1.4 KAVAGait: clinical gait analysis

KAVAGait [76] is a knowledge-assisted VA system for clinical gait analysis (see Figure 6) that supports analysts during diagnosis and clinical decision making. Users can load patient gait data containing *ground reaction forces* (GRF) measurements. These collected GRF data are visualized as wave forms in the center of the interface, representing a separated view for the left (red) and the right (blue) foot as well as providing a combined visualization. Additionally, 16 *spatio-temporal parameters* (STP) (e.g., step time, stance time, cadence) are calculated, visualized, and used for automated patient comparison and categorization.

Since one primary goal during clinical gait analysis is to assess whether a recorded gait measurement displays *normal gait* behavior or if not, which specific *gait abnormality* are present. Thus, the system's internal *explicit knowledge store* (EKS) contains several

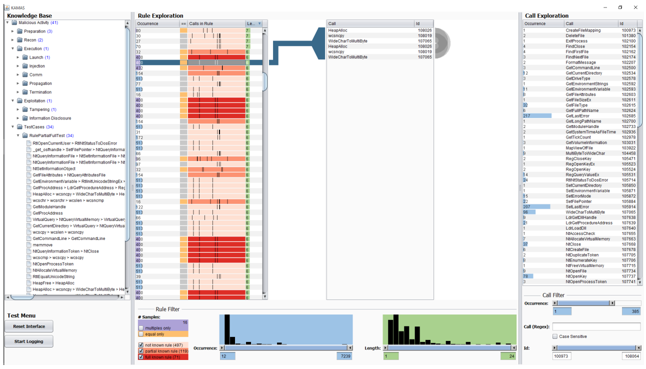

Figure 5: KAMAS, a knowledge-assisted malware analysis system [75], supporting IT-security experts during behavior-based malware analysis.

categories of gait abnormalities (e.g., knee, hip, ankle) as well as a category including healthy gait pattern data. Each category is defined by a set of parameter ranges $[min, max]$ of the 16 calculated STPs. All EKS entries are used for analysis and comparison by default. However, analysts can apply their expertise (tacit knowledge) as specification $\boxed{K^\tau} \rightarrow E \rightarrow \boxed{S}$, to filter entries by patient data (e.g., age, height, weight).

Automated data analysis of newly loaded patient data is provided for categories (e.g., automatically calculated category matching) influencing the systems specification: $\{\boxed{D}, \boxed{K^\varepsilon}, \boxed{S}\} \rightarrow A \rightarrow \boxed{S}$. The EKS stores **explicit knowledge** and the automated data analysis methods are strongly intertwined with the visual data analysis system in KAVAGait. Thus, the combined analysis and visualization pipeline consists of the following process chain, and supports the analysts while interactive data exploration $\{\boxed{D}, \{\boxed{D}, \boxed{K^\varepsilon}, \boxed{S}\} \rightarrow A \rightarrow \boxed{S}\} \rightarrow V$. Based on the visualization, the generated image is perceived by the analyst, gaining **tacit knowledge** $V \xrightarrow{I} P \rightarrow \boxed{K^\tau}$, which also influences the analysts perception $\boxed{K^\tau} \rightarrow P$. As data **exploration and analysis** is an iterative process, the analyst gains further tacit knowledge based on the adjusted visualization and driven by the specification. To generate explicit knowledge, the analyst can include the STPs of analyzed patients based on his/her clinical decisions to the EKS, which can be described as the extraction of tacit knowledge $\boxed{K^\tau} \rightarrow X \rightarrow \boxed{K^\varepsilon}$.

Moreover, KAVAGait provides the ability to **interactively explore and adjust the systems EKS**, whereby the explicit knowledge can be visualized in a separated view $\boxed{K^\varepsilon} \rightarrow V$. Two different options (one for a single patient and one for a category) are provided in KAVAGait for the adjustment of the stored explicit knowledge by the analysts' tacit knowledge. $\boxed{K^\varepsilon} \rightarrow V \xrightarrow{I} P \rightarrow \boxed{K^\tau} \rightarrow X \rightarrow \boxed{K^\varepsilon}$.

4.2 Assessing Costs and Profits of Explicit Knowledge

Second, the knowledge-assisted VA model can be a framework to compare different design alternatives. As specified by van Wijk [74] we are assuming that a community of n homogeneous users are using the visualization V to visualize a dataset m times. Therefore, each user needs j exploration steps per session and a time t. Additionally, in "the real world, the user community will often be highly varied, with different K_0^τ's and also with different aims" [74]. Thus, the four types of costs: *Initial Development Costs* $C_i(S_0)$; *Initial Costs per User* $C_u(S_0)$; *Initial Costs per Session* $C_s(S_0)$ and *Perception and Exploration Costs* C_e [74] can be extended with the generation of explicit knowledge K^ε based on l knowledge generation steps. This *Knowledge Extraction and Computerization Costs* C_k are related to the users' tacit knowledge extraction, the knowledge generation by automated data analysis methods, and the combination of both. Based on these five cost elements, the total costs C can be described as their sum. Additionally, the knowledge gain G can be described by the generated tacit knowledge ΔK^τ by the user as well as the extracted explicit knowledge ΔK^ε added to the system per session, which have to be multiplied by the total number of sessions. Based on the calculated costs C and the knowledge gain G, the total profit F of of the system can be described by $F = G - C$ according to the description of van Wijk [74].

Generally, this description tells us that a successful knowledge-assisted VA system is used by many users, gaining high values from knowledge and extracting it to the system without spending time and money on hardware and training [74]. The more tacit knowledge users gain during data exploration, the more explicit knowledge can be included into the system. The user gets the ability to use explicit knowledge generated by herself, by others, and by automated analysis methods to achieve her goals. Thus, VA is not only improved but also accelerated. Additionally, by sharing knowledge in explicit form, users get the opportunity to learn from others, to improve and gain new insights.

From the perspective of interaction costs (approximately a combination of C_e, $C_u(S_0)$, $C_s(S_0)$), which are described by Lam [49] as "less is more", can be optimized by reducing the effort of execution and evaluation. Thereby, the knowledge-assisted VA process moves parts of the specification effort from the human via E to machine via A. Additionally, automated analysis methods are supporting the user by analyzing the data based on $\boxed{S}$ and $\boxed{K^\varepsilon}$. Thus, the analyst has the ability to gain new tacit knowledge $\boxed{K^\tau}$ which can be extracted as $\boxed{K^\varepsilon}$ to adjust $\boxed{S}$ and A.

Chen and Golan [20] suggest that the most generic cost is energy for both the computer (e.g., run an algorithm, create a visualization) and the human (e.g., read data, view visualization, decision making). A measurement of the computers' energy consumption is common practice, but the measurement of the users' activities is mostly not feasible [20]. Therefore, time t can be a point for the measurement as well as the amount of performed exploration steps j and knowledge generation steps l. Additionally, Crouser et al. [26] state that a model currently cannot elaborate how much a user is doing, its only possible to measure how often the human is working. Tam et al. [70] introduce an information-theoretic model to analyze both the machine and the human contribution to the VA process, in particular for a classification task. Kijmongkolchai et al. [47] propose a methodology to empirically measure human's soft knowledge, confirming that it can enhance the cost-benefit ratio of a visualization process.

Our novel knowledge-assisted VA model (see Figure 1), enables the identification of the contribution to knowledge generation by

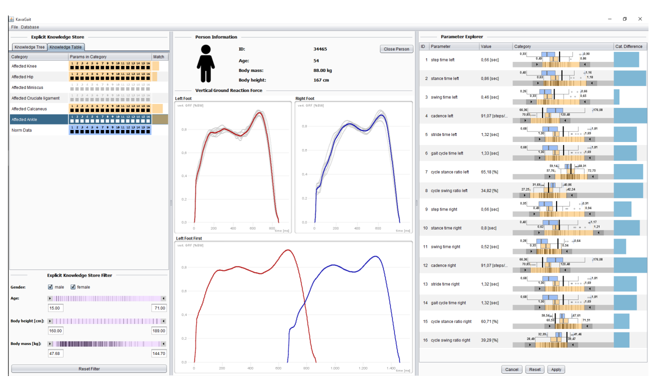

Figure 6: KAVAGait, a knowledge-assisted clinical gait analysis system [76], supporting analysts during clinical decision making.

the human and by the machine, through two distinct processes: the human perception P and the automated analysis A. Analogously, it distinguishes the contributions to the specification $\boxed{S}$ by the human (through the exploration process E) and by the machine (through the guidance process G). In other words, by identifying profits and costs on both the human and the machine side, our model provides the basis for evaluating the performance of knowledge-assisted VA environments.

4.3 Inspiring Innovative Approaches and Research

The generality of the model has been demonstrated above by applying it to the description of several different examples of knowledge-assisted VA. Its generality and its operational nature can eventually enable the description of future knowledge-assisted systems.

In this section, in particular, we discuss how the model can also work as a conceptual tool in VA design, in order to conceive new processes and scenarios involving explicit knowledge, and, therefore, it can inspire future research directions. By reasoning about the different dimensions (knowledge processes, types, and sources) and the combination thereof in existing approaches, new possibilities can be generated by analogy, opposition, transposition, or symmetry, not only by introducing new artifacts and processes, but also considering the extension of existing processes to different knowledge types and sources. In the following, we demonstrate several examples of this generation; we start with assessment of existing scenarios and discuss potentiality and plausibility of new ones.

From prior research on theory we know how different types of tacit knowledge are used on the human side (see Section 2.1): operational knowledge supports mainly exploration while domain knowledge supports cognition and perception. From our model and its application to several examples (see Section 4.1), we observed that the use of explicit knowledge on the computer side mimics tacit knowledge: operational knowledge is generally exploited in processes involving specification, while domain knowledge in processes involving data. However, some systems also use domain knowledge for guidance and specification as for example systems mapping domain ontologies (representing declarative domain knowledge) to visual representation taxonomies in order to suggest adequate visual encodings [16, 36] or systems exploiting declarative domain knowledge to assist and guide interaction [33]. Additionally, domain ontologies can be mapped to analytical algorithms taxonomies – in doing so, the choice of analytical algorithms and their parameters can be guided in a similar way to their visualization counterpart.

Moreover, since domain knowledge can be exploited for specification processes it would also be possible to use operational knowledge not only to control but also to augment data and data analysis. We can imagine processes to extract domain knowledge from operational knowledge, in analogy to recommender systems – if many users put their focus on two data items (or data classes, or data sets) via interaction, it is likely that they are related. The system can store this information in its knowledge base, and utilize it for analysis and visualization.

Approaches that try to infer operational knowledge by mining user interactions generally examine the exploration process, which operates across the human-computer interface; however, also the perception process operates across the human-computer interface and can be automatically observed to assist interaction (e.g., interaction by eye-gaze input [42]). It is reasonable to envision an integration of explicit knowledge within this kind of approaches, leading for example to a possible semantic gaze-based interaction.

We have defined explicit and tacit knowledge from a human-centric perspective, assuming that internal knowledge retained by humans is not necessarily accessible and easily transferable, while knowledge processed and stored by computers is both, accessible and transferable. This perspective is coherent with recent advances in progressive VA [68] and the human-is-the-loop paradigm [29]: results of data analysis, inner states of algorithms, and their parameters should always be accessible, understandable, and steerable by the user. However, by observing the symmetry in our model between the human space and the machine space, we can posit the existence of tacit knowledge inside algorithms; in other words, we posit a knowledge resulting from computer-simulated cognitive processes which are not easily accessible and interpretable by the human counterpart. Indeed, this might be the case if we integrate deep learning methods as the analytical component of a VA process, since these methods not always permit knowledge representation which facilitates the human in the analytic discourse. Therefore, further research would be needed to ensure provenance, awareness, and trust in these scenarios.

4.4 Discussion

As Beaudouin-Lafon pointed out, a good theoretical model needs to "strike a balance between generality (for descriptive power), concreteness (for evaluative power) and openness (for generative power)" [6, p. 17], which are contradicting goals. Our new Knowledge-assisted VA Model represents a high-level system blueprint, which can be used for a generalized system description from the viewpoint of the components to be used, the included processes and their connections. However, the model does not provide the system architecture at the level of detail which is required directly for the implementation (e.g., design patterns, algorithms, data structures).

Model Comparison: Since this model focuses on high-level systems architecture, a limitation is the possibility of describing the cognitive processes, perception and tacit knowledge generation of users. For that, other established models like the Knowledge Generation Model for VA by Sacha et al. [62] might be used.

The three loops of the Knowledge Generation Model for VA for exploration, verification, and knowledge generation are tightly intertwined where lower-level loops are directed by higher-level loops. Each of these three loops can be reconstructed and described with the new 'Knowledge-assisted VA Model'. To demonstrate this, we are describing the Knowledge Generation Loop as example: The entire verification process is driven by the analyst's tacit knowledge. There are several types of knowledge and we can distinguish between two general phases of externalizing (explicit knowledge $\boxed{K^\varepsilon}$) and internalizing knowledge (tacit knowledge $\boxed{K^\tau}$). Hypothesis and assumptions about the data are defined and formed based on tacit knowledge and trusted and verified insights are internalized as new tacit knowledge. Moreover, VA allows the analyst to provide feedback to the system in order to incorporate the analysts knowledge into the entire process. This can be also achieved by extracting the analysts tacit knowledge to the system where it is made available as explicit knowledge in a computerized form. Seen from the view of the Knowledge-assisted VA Model, the analysts tacit knowledge can be externalized and included into the system as explicit knowledge: $\boxed{K^\tau} \rightarrow X \rightarrow \boxed{K^\varepsilon}$. This explicit knowledge is then included into the VA process to influence the automated data analysis methods and/or to change the systems specification: $\boxed{K^\varepsilon} \rightarrow \{ A, A \rightarrow \boxed{S} \}$. Additionally, based on the experts tacit knowledge, the systems specification can be manipulated directly. Depending on specification, the automated data analysis methods and the visualization are adjusted: $\boxed{K^\tau} \rightarrow E \rightarrow \boxed{S} \rightarrow \{ A, V \}$. Additionally, it is also possible to perform indirect adjustments for the explicit knowledge and the data: $\boxed{S} \rightarrow A \rightarrow \{ \boxed{D}, \boxed{K^\varepsilon} \}$.

Model Limitations: As demonstrated above, all three loops included in the model by Sacha et al. [62], can also be recreated by the new Knowledge-assisted VA Model. In general, the Knowledge Generation Model for VA fits better for the description of the performed operations by the user. In contrast, the new Knowledge-assisted VA Model can be used to describe the systems characteristics. Based on a combination of both models, the designer gets the ability to describe the user processes at a more detailed level with respect to the included components and processes to generate a detailed system abstraction. Moreover, the new Knowledge-assisted VA Model does not detail the way how explicit knowledge and analysis data are collected, prepared, stored or made available. Last but not least, the model provides an theoretical approach to calculate the costs, the knowledge gain and the profit of knowledge-assisted VA systems, but it does not provide any procedure to measure and quantify the quality of the integrated explicit knowledge or to prevent the analyst in terms of misleading knowledge.

5 Conclusion & Future Work

The main contribution of this work is the extension of the theoretical underpinnings of VA in order to incorporate the function and role of tacit and explicit knowledge in the analytical reasoning process. We propose a novel conceptual model that generalizes existing approaches of Knowledge-assisted VA. It is based on the well-known visualization model of van Wijk [74] and allows for modeling a broad range of analytics systems (both with and without explicit knowledge as well as automated data analysis). Hence, it connects seamlessly to existing theoretical foundations while extending their descriptive, evaluative, and generative power. The new model contains all components, processes, and connections needed in a Knowledge-assisted VA system, i.e. 1) tacit knowledge extraction; 2) automated data analysis methods; 3) explicit knowledge based specification; 4) explicit knowledge visualization; and 5) tacit knowledge generation.

In the paper, we illustrated the possibilities of explicit knowledge integration and extraction, the integration of automated data analysis methods as well as the combination of both. This supports data exploration, analysis, and gain of tacit knowledge as well as the extraction of knowledge and its sharing with other users.

We demonstrated the utility of the model by showing how it can be used to 1) describe the characteristics of a broad range of existing approaches; 2) evaluate the costs and benefits of knowledge-assisted processes and systems; and 3) inspire and enable the design of innovative approaches as a high-level system blueprint.

As this work represents an early step in this area, a number of opportunities for future research arise. One issue of major necessity are novel evaluation methods that can measure knowledge flows to assess the effectiveness of VA environments. Such methods can be based on explicit knowledge as conceptualized in our model. For example, the nested workflow model [31] points in this direction, enabling the description of VA processes also in terms of data and knowledge flows. Further areas of future research are validation methods for extracted explicit knowledge, extracting knowledge indirectly via user interactions, or more specific support for collaboration and multi-user systems.

Acknowledgments

The authors wish to thank Heidrun Schumann, Christian Tominski, Daniel A. Keim, and Jarke J. van Wijk for constructive discussions. This work was supported in part by the Austrian Science Fund (FWF) via the KAVA-Time project (P25489-N23) and the VisOnFire project (P27975-NBL), by the Austrian Research Promotion Agency (FFG) via the Devisor project (850695), and by the Austrian Federal Ministry of Science, Research, and Economy in the Laura Bassi Centres of Excellence initiative (CVAST, 840262).

References

[1] R. Ackoff. From data to wisdom. *Journal of Applied Systems Analysis*, 16:3–9, 1989.

[2] W. Aigner and S. Miksch. CareVis: Integrated visualization of computerized protocols and temporal patient data. *Artificial Intelligence in Medicine*, 37(3):203–218, 2006. doi: 10.1016/j.artmed.2006.04.002

[3] W. Aigner, A. Rind, and S. Hoffmann. Comparative evaluation of an interactive time-series visualization that combines quantitative data with qualitative abstractions. *Computer Graphics Forum*, 31(3pt2):995–1004, 2012. doi: 10.1111/j.1467-8659.2012.03092.x

[4] R. Bade, S. Schlechtweg, and S. Miksch. Connecting time-oriented data and information to a coherent interactive visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '04, pp. 105–112. ACM, New York, NY, USA, 2004. doi: 10.1145/985692.985706

[5] R. C. Basole, H. Park, M. Gupta, M. L. Braunstein, D. H. Chau, and M. Thompson. A visual analytics approach to understanding care process variation and conformance. In *Proceedings of the 2015 Workshop on Visual Analytics in Healthcare*, VAHC '15, pp. 6:1–6:8. ACM, New York, NY, USA, 2015. doi: 10.1145/2836034.2836040

[6] M. Beaudouin-Lafon. Designing interaction, not interfaces. In *Proc. Working Conf. Advanced Visual Interfaces, AVI*, pp. 15–22. ACM, 2004. doi: 10.1145/989863.989865

[7] B. B. Bederson and B. Shneiderman. Theories for understanding information visualization. In *The Craft of Information Visualization: Readings and Reflections*, pp. 349–351. Morgan Kaufmann, San Francisco, CA, 2003.

[8] E. Bertini and D. Lalanne. Surveying the complementary role of automatic data analysis and visualization in knowledge discovery. In *Proc. ACM Int. Conf. on Knowl. Discovery and Data Min. (SIGKDD)*, pp. 12–20, 2009. doi: 10.1145/1562849.1562851

[9] P. Bodesinsky, P. Federico, and S. Miksch. Visual analysis of compliance with clinical guidelines. In *Proceedings of the 13th International Conference on Knowledge Management and Knowledge Technologies*, i-Know '13, pp. 12:1–12:8. ACM, New York, NY, USA, 2013. doi: 10.1145/2494188.2494202

[10] F. Bouali, A. Guettala, and G. Venturini. VizAssist: an interactive user assistant for visual data mining. *The Visual Computer*, 32(11):1447–1463, 2016. doi: 10.1007/s00371-015-1132-9

[11] N. Boukhelifa, W. Cancino, A. Bezerianos, and E. Lutton. Evolutionary visual exploration: Evaluation with expert users. *Computer Graphics Forum*, 32(3pt1):31–40, 2013. doi: 10.1111/cgf.12090

[12] B. Broeksema, T. Baudel, A. Telea, and P. Crisafulli. Decision exploration lab: A visual analytics solution for decision management. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):1972–1981, Dec 2013. doi: 10.1109/TVCG.2013.146

[13] E. T. Brown, A. Ottley, H. Zhao, Q. Lin, R. Souvenir, A. Endert, and R. Chang. Finding Waldo: Learning about users from their interactions. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1663–1672, Dec 2014. doi: 10.1109/TVCG.2014.2346575

[14] S. K. Card and J. D. Mackinlay. The structure of the information visualization design space. In *IEEE Symposium on Information Visualization (InfoVis)*, pp. 92–99, 1997. doi: 10.1109/INFVIS.1997.636792

[15] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in Information Visualisation. Using Vision to Think*. Morgan Kaufman Publ Inc, San Francisco, CA, Aug. 1999.

[16] S. Carpendale, M. Chen, D. Evanko, N. Gehlenborg, C. Görg, L. Hunter, F. Rowland, M. A. Storey, and H. Strobelt. Ontologies in biological data visualization. *IEEE Computer Graphics and Applications*, 34(2):8–15, Mar 2014. doi: 10.1109/MCG.2014.33

[17] D. Ceneda, T. Gschwandtner, T. May, S. Miksch, H. J. Schulz, M. Streit, and C. Tominski. Characterizing guidance in visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):111–120, Jan 2017. doi: 10.1109/TVCG.2016.2598468

[18] C. Chen. Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16, July 2005. doi: 10.1109/MCG.2005.91

[19] M. Chen, D. Ebert, H. Hagen, R. Laramee, R. Van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. *IEEE Computer Graphics and Applications*, 29(1):12–19, Jan. 2009. doi: 10.1109/MCG.2009.6

[20] M. Chen and A. Golan. What may visualization processes optimize? *IEEE Transactions on Visualization and Computer Graphics*, 22(12):2619–2632, Dec 2016. doi: 10.1109/TVCG.2015.2513410

[21] M. Chen and H. Hagen. Guest editors' introduction: Knowledge-assisted visualization. *IEEE Computer Graphics and Applications*, 30(1):15–16, 2010.

[22] E. H. Chi. A taxonomy of visualization techniques using the data state

reference model. In *IEEE Symposium on Information Visualization (InfoVis)*, pp. 69–75, 2000. doi: 10.1109/INFVIS.2000.885092

[23] E. H. Chi. Expressiveness of the data flow and data state models in visualization systems. In *Proc. of the Working Conference on Advanced Visual Interfaces*, AVI '02, pp. 375–378. ACM, New York, NY, USA, 2002. doi: 10.1145/1556262.1556327

[24] E. H.-H. Chi and J. Riedl. An operator interaction framework for visualization systems. In *IEEE Symposium on Information Visualization (InfoVis)*, pp. 63–70, Oct. 1998. doi: 10.1109/INFVIS.1998.729560

[25] D. H. S. Chung, M. L. Parry, I. W. Griffiths, R. S. Laramee, R. Bown, P. A. Legg, and M. Chen. Knowledge-assisted ranking: A visual analytic application for sports event data. *IEEE Computer Graphics and Applications*, 36(3):72–82, May 2016. doi: 10.1109/MCG.2015.25

[26] R. J. Crouser, L. Franklin, A. Endert, and K. Cook. Toward theoretical techniques for measuring the use of human effort in visual analytic systems. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):121–130, Jan 2017. doi: 10.1109/TVCG.2016.2598460

[27] F. Dabek and J. J. Caban. 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, Jan 2017. doi: 10.1109/TVCG.2016.2598471

[28] A. Endert. Semantic interaction for visual analytics: Toward coupling cognition and computation. *IEEE Computer Graphics and Applications*, 34(4):8–15, July 2014. doi: 10.1109/MCG.2014.73

[29] A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews. The human is the loop: New directions for visual analytics. *J. Intell. Inf. Syst.*, 43(3):411–435, Dec. 2014. doi: 10.1007/s10844-014-0304-9

[30] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth. From data mining to knowledge discovery in databases. *AI magazine*, 17(3):37, 1996.

[31] P. Federico, A. Amor-Amorós, and S. Miksch. A nested workflow model for visual analytics design and validation. In *Proc. of the Workshop on Beyond Time And Errors (BELIV)*, BELIV '16, pp. 104–111. ACM, New York, NY, USA, 2016. doi: 10.1145/2993901.2993915

[32] P. Federico, S. Hoffmann, A. Rind, W. Aigner, and S. Miksch. Qualizon graphs: Space-efficient time-series visualization with qualitative abstractions. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces*, AVI '14, pp. 273–280. ACM, New York, NY, USA, 2014. doi: 10.1145/2598153.2598172

[33] P. Federico, J. Unger, A. Amor-Amorós, L. Sacchi, D. Klimov, and S. Miksch. Gnaeus: Utilizing clinical guidelines for knowledge-assisted visualisation of EHR cohorts. In E. Bertini and J. C. Roberts, eds., *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association, 2015. doi: 10.2312/eurova.20151108

[34] S. Garg, J. E. Nam, K. Padalkar, S. Laue, W. Saleem, J. Giesen, and K. Mueller. KAV-DB: Towards a framework for the capture and retrieval of visualization knowledge over the web. In *Proceedings of the Schloss Dagstuhl Scientific Visualization Workshop 33(5) (SciVis)*, pp. 607–615, 2010.

[35] S. Garg, J. E. Nam, I. V. Ramakrishnan, and K. Mueller. Model-driven visual analytics. In *2008 IEEE Symposium on Visual Analytics Science and Technology*, pp. 19–26, Oct 2008. doi: 10.1109/VAST.2008.4677352

[36] O. Gilson, N. Silva, P. Grant, and M. Chen. From web data to visualization via ontology mapping. *Computer Graphics Forum*, 27(3):959–966, 2008. doi: 10.1111/j.1467-8659.2008.01230.x

[37] T. M. Green, W. Ribarsky, and B. Fisher. Building and applying a human cognition model for visual analytics. *Information Visualization*, 8(1):1–13, 2009. doi: 10.1057/ivs.2008.28

[38] T. Gschwandtner, W. Aigner, K. Kaiser, S. Miksch, and A. Seyfang. CareCruiser: Exploring and visualizing plans, events, and effects interactively. In *2011 IEEE Pacific Visualization Symposium*, pp. 43–50, March 2011. doi: 10.1109/PACIFICVIS.2011.5742371

[39] J. Han, X. Hu, and N. Cercone. A visualization model of interactive knowledge discovery systems and its implementations. *Information Visualization*, 2(2):105–125, June 2003. doi: 10.1057/palgrave.ivs.9500045

[40] D. J. Hand. Intelligent data analysis: Issues and opportunities. *Intell. Data Anal.*, 2:67–79, 1998. doi: 10.1016/S1088-467X(99)80001-8

[41] E. Horvitz. Principles of mixed-initiative user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pp. 159–166. ACM, New York, NY, USA, 1999. doi: 10.1145/302979.303030

[42] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey. Human-computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6):1527–1534, Nov 1989. doi: 10.1109/21.44068

[43] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *IVS*, 10(4):310–326, 2011. doi: 10.1177/1473871611412817

[44] B. Kamsu-Foguem, G. Tchuenté-Foguem, L. Allart, Y. Zennir, C. Vilhelm, H. Mehdaoui, D. Zitouni, H. Hubert, M. Lemdani, and P. Ravaux. User-centered visual analysis using a hybrid reasoning architecture for intensive care units. *Decision Support Systems*, 54(1):496–509, 2012. doi: 10.1016/j.dss.2012.06.009

[45] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, eds. *Mastering the information age: solving problems with visual analytics*. Eurographics Association, Goslar, 2010.

[46] D. A. Keim, F. Mansmann, J. Schneidewind, J. Thomas, and H. Ziegler. Visual Analytics: Scope and challenges. In S. J. Simoff, M. H. Böhlen, and A. Mazeika, eds., *Visual Data Mining*, LNCS 4404, pp. 76–90. Springer, Berlin, 2008. doi: 10.1007/978-3-540-71080-6_6

[47] N. Kijmongkolchai, A. Abdul-Rahman, and M. Chen. Empirically measuring soft knowledge in visualization. *Computer Graphics Forum*, 36(3):073–085, 2017. doi: 10.1111/cgf.13169

[48] D. Klimov, Y. Shahar, and M. Taieb-Maimon. Intelligent visualization and exploration of time-oriented data of multiple patients. *Artificial Intelligence in Medicine*, 49(1):11–31, 2010. doi: 10.1016/j.artmed.2010.02.001

[49] H. Lam. A framework of interaction costs in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1149–1156, Nov 2008. doi: 10.1109/TVCG.2008.109

[50] T. Lammarsch, W. Aigner, A. Bertone, S. Miksch, and A. Rind. Towards a concept how the structure of time can support the visual analytics process. In S. Miksch and G. Santucci, eds., *Proc. of the EuroVis Workshop on Visual Analytic (EuroVA)*, pp. 9–12. Eurographics, Eurographics, 2011. doi: 10.2312/PE/EuroVAST/EuroVA11/009-012

[51] F. T. Marchese and E. Banissi, eds. *Knowledge Visualization Currents*. Springer, London, 2013.

[52] G. Mistelbauer, H. Bouzari, R. Schernthaner, I. Baclija, A. Köchl, S. Bruckner, M. Sramek, and M. E. Gröller. Smart Super Views – a knowledge-assisted interface for medical visualization. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 163–172, Oct 2012. doi: 10.1109/VAST.2012.6400555

[53] M. Mokhtari, E. Boivin, D. Laurendeau, and M. Girardin. Visual tools for dynamic analysis of complex situations. In *2010 IEEE Symposium on Visual Analytics Science and Technology*, pp. 241–242, Oct 2010. doi: 10.1109/VAST.2010.5654451

[54] A. Motamedi, A. Hammad, and Y. Asen. Knowledge-assisted BIM-based visual analytics for failure root cause detection in facilities management. *Automation in Construction*, 43:73–83, 2014. doi: 10.1016/j.autcon.2014.03.012

[55] J. E. Nam, M. Maurer, and K. Mueller. A high-dimensional feature clustering approach to support knowledge-assisted visualization. *Computers & Graphics*, 33(5):607–615, 2009. doi: 10.1016/j.cag.2009.06.006

[56] P. Perner. Intelligent data analysis in medicine—recent advances. *Artif. Intell. Med.*, 37(1):1–5, May 2006. doi: 10.1016/j.artmed.2005.10.003

[57] W. A. Pike, J. Stasko, R. Chang, and T. A. O'Connell. The science of interaction. *Information Visualization*, 8(4):263–274, 2009. doi: 10.1057/ivs.2009.22

[58] P. Pirolli and S. Card. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of the International Conference on Intelligence Analysis*, vol. 5, pp. 2–4, 2005.

[59] W. Ribarsky and B. Fisher. The human-computer system: Towards an operational model for problem solving. In *Hawaii Int. Conf. on System Sciences (HICSS)*, pp. 1446–1455, Jan. 2016. doi: 10.1109/HICSS.2016.183

[60] J. Rowley. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2):163–180, 2007. doi: 10.1177/0165551506070706

[61] S. Rubrichi, C. Rognoni, L. Sacchi, E. Parimbelli, C. Napolitano, A. Mazzanti, and S. Quaglini. Graphical representation of life paths to better convey results of decision models to patients. *Medical Decision Making*, 35(3):398–402, 2015. doi: 10.1177/0272989X14565822

[62] D. Sacha, A. Stoffel, F. Stoffel, B. C. Kwon, G. Ellis, and D. 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

[63] B. Saket, H. Kim, E. T. Brown, and A. Endert. Visualization by demonstration: An interaction paradigm for visual data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):331–340, Jan 2017. doi: 10.1109/TVCG.2016.2598839

[64] Y. Shahar. A framework for knowledge-based temporal abstraction. *Artificial Intelligence*, 90(12):79–133, 1997. doi: 10.1016/S0004-3702(96)00025-2

[65] Y. B. Shrinivasan and J. J. van Wijk. Vispad: Integrating visualization, navigation and synthesis. In *2007 IEEE Symposium on Visual Analytics Science and Technology*, pp. 209–210, Oct 2007. doi: 10.1109/VAST.2007.4389021

[66] A. M. Smith, W. Xu, Y. Sun, J. R. Faeder, and G. E. Marai. RuleBender: integrated modeling, simulation and visualization for rule-based intracellular biochemistry. *BMC Bioinf.*, 13(8):S3, 2012. doi: 10.1186/1471-2105-13-S8-S3

[67] M. Steiger, T. May, J. Davey, and J. Kohlhammer. Visual analysis of expert systems for smart grid monitoring. In M. Pohl and H. Schumann, eds., *EuroVis Workshop on Visual Analytics*. The Eurographics Association, 2013. doi: 10.2312/PE.EuroVAST.EuroVA13.043-047

[68] C. D. Stolper, A. Perer, and D. Gotz. Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, Dec 2014. doi: 10.1109/TVCG.2014.2346574

[69] E. Swing. Prajna: Adding automated reasoning to the visual- analysis process. *IEEE Computer Graphics and Applications*, 30(1):50–58, Jan 2010. doi: 10.1109/MCG.2010.15

[70] G. K. L. Tam, V. Kothari, and M. Chen. An analysis of machine- and human-analytics in classification. *TVCG*, 23(1):71–80, Jan 2017. doi: 10.1109/TVCG.2016.2598829

[71] S.-O. Tergan and T. Keller, eds. *Knowledge and Information Visualization*, vol. 3426 of *LNCS*. Springer, Berlin, 2005.

[72] J. J. Thomas and K. A. Cook. *Illuminating the path: the research and development agenda for visual analytics*. IEEE Computer Society, 2005.

[73] J. Vallet, B. Pinaud, and G. Melançon. Studying propagation dynamics in networks through rule-based modeling. In *IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 281–282, Oct 2014. doi: 10.1109/VAST.2014.7042530

[74] J. J. van Wijk. The value of visualization. In *Proc. IEEE Visualization (VIS 05)*, pp. 79–86, 2005. doi: 10.1109/VISUAL.2005.1532781

[75] M. Wagner, A. Rind, N. Thür, and W. Aigner. A knowledge-assisted visual malware analysis system: design, validation, and reflection of KAMAS. *Computers & Security*, 67:1–15, 2017. doi: 10.1016/j.cose.2017.02.003

[76] M. Wagner, D. Slijepcevic, B. Horsak, A. Rind, M. Zeppelzauer, and W. Aigner. KAVAGait: Knowledge-assisted visual analytics for clinical gait analysis. *arXiv:1707.06105 [cs.HC]*, July 2017.

[77] X. Wang, W. Dou, S.-E. Chen, W. Ribarsky, and R. Chang. An interactive visual analytics system for bridge management. *Computer Graphics Forum*, 29(3):1033–1042, 2010. doi: 10.1111/j.1467-8659.2009.01708.x

[78] X. Wang, D. H. Jeong, W. Dou, S.-W. Lee, W. Ribarsky, and R. Chang. Defining and applying knowledge conversion processes to a visual analytics system. *Computers & Graphics*, 33(5):616–623, Oct. 2009. doi: 10.1016/j.cag.2009.06.004

[79] M. Workman, M. F. Lesser, and J. Kim. An exploratory study of cognitive load in diagnosing patient conditions. *Int. J. Qual. Health Care*, 19(3):127, 2007. doi: 10.1093/intqhc/mzm007

[80] M. Zeleny. Management support systems: towards integrated knowledge management. *Human systems management*, 7(1):59–70, 1987.

[81] B. Zupan, J. H. Holmes, and R. Bellazzi. Knowledge-based data analysis and interpretation. *Artif. Intell. Med.*, 37(3):163–165, 2006. doi: 10.1016/j.artmed.2006.03.001