A Nested Workflow Model for Visual Analytics Design and Validation

Paolo Federico
Inst. of Software Technology
and Interactive Systems
TU Wien, Austria
federico@ifs.tuwien.ac.at

Albert Amor-Amorós
Inst. of Software Technology
and Interactive Systems
TU Wien, Austria
amor@ifs.tuwien.ac.at

Silvia Miksch
Inst. of Software Technology
and Interactive Systems
TU Wien, Austria
miksch@tuwien.ac.at

ABSTRACT

Characterizing the problem domain and understanding users' practices and processes are recognized as important steps in order to design and validate visualization, but are often disregarded in practice, also because of their complexity. We introduce the nested workflow model for design and validation of visual analytics, aimed at providing designers with a powerful and expressive modelling tool. This model enables the description of visual analytics processes, at different design levels, in terms of tasks, data, and users, including complex workflow patterns, data and knowledge flows, and collaboration between users. We discuss its application to two visual analytics projects, demonstrating its usefulness for their design and validation.

CCS Concepts

•**Human-centered computing → Visualization design and evaluation methods;**

Keywords

visualization, design, validation, nested model, workflow, processes, problem domain characterization, visual analytics

1. INTRODUCTION

The nested model by Munzner [27] provides a robust and sound methodological framework for design and validation of visualization. It features four layers: (i) characterization of the problem domain, (ii) abstraction of tasks and data from a domain-specific to a domain-independent form, (iii) design of visual encodings and interaction techniques addressing the tasks and data, and (iv) creation of algorithms to implement the visual encodings and interaction techniques.

These level are nested, thus the output from an upstream level above is input to the downstream level below. Since each level faces its own threats to validity, Munzner's model helps visualization designers with positioning their contribution in the right level, choosing the adequate validation strategy for that level, and understanding assumptions and implications with regards to other levels. In particular, the nested model highlights the importance of the problem characterization level, which is usually underrepresented in the visualization literature.

The nested blocks and guidelines model (NBGM) [25] extends the four-level nested model, and complements its methodological grounding with operational directions. In particular, the NBGM adds a finer-grained structure within each level by means of blocks, which can be either identified or designed. The blocks of the domain level are identified, since they represent the understanding of domain situations by designers; conversely, the blocks of other levels are designed, since they represent the outcomes of design choices taken within each level. Within-level guidelines enable comparison between blocks at the same level, while between-level guidelines represent mapping rules between blocks at adjacent levels. As a result, the NBGM helps designers with representing and discussing design rationales.

Grounded within the four-level methodological framework, we propose a nested workflow model for design and validation of visual analytics. It is inspired by the NBGM, and features nested blocks as components of the analytical workflow. Our model is aimed at supporting designers with the characterization of the problem domain, by capturing users' practices and processes with respect to three aspects: tasks, data, and users. The formulation of the proposed model has been mainly driven by a reflection on our experiences with the problem domain level, but its nested structure presents implications for other levels as well.

The main contribution of our work is the introduction of the nested workflow model. In the following, we recall workflows concepts and terminology; then we introduce our model and discuss how it addresses the problem domain characterization and the other nested levels, with particular consideration of tasks, data, and users; finally, we demonstrate its application to two visual analytics research projects: identification of unexploded ordnance risks and shared decision-making about medical treatment plans between cardiologists and patients.

2. WORKFLOWS: CONCEPTS AND TERMINOLOGY

Our nested workflow model for visual analytics leverages concepts and terms from business process modelling and workflow management, which we want to mention before describing our model. In the context of workflow management, a *workflow* is defined as "the automation of a business pro-

BELIV '16, October 24 2016, Baltimore, MD, USA

DOI: http://dx.doi.org/10.1145/2993901.2993915

cess, in whole or part, during which documents, information or tasks are passed from one participant to another for action (activities), according to a set of procedural rules" [47, page 8]. In the context of business process modelling, the term *workflow* can be also used as a synonym for *process* [41]; we use this acceptation, since we are interested in the modelling of processes but not in their full automation (see also Section 3). A *workflow* consists of a number of tasks. A *task* is a unit of work that is carried out as a single whole by one or a group of actors (also known as participants or resources), and an *actor* is either a person or a machine performing specific tasks [41].

The execution of tasks can be arranged according to a number of structural *patterns* [40]: e.g., *sequence* (the execution of one task starts after the completion of the previous one), *parallelism* (several tasks can be executed in parallel), *synchronization* (several tasks can be executed in parallel, but they all need to be completed before another one can start), *selection* or *choice* (the workflow splits in several branches, and only one branch is executed), *iteration* (a task or a set of tasks can be repeated).

In order to perform tasks, humans need *knowledge* [40], which can be either *tacit* (prior knowledge acquired by experience and stored in their minds) or *explicit* (formalized and externalized knowledge, that can be accessed and retrieved from external sources).

Workflows can be divided into business workflows and scientific workflows, with two relevant differences in this context: the interaction with participants and the flow paradigm. As for the interaction with participants, the tasks of business worflows can be generally executed by humans or machines, while scientific workflow consist of data processing tasks that are generally performed by machines, under the control of humans [22]. As for the flow paradigm, the order of tasks in business workflows is determined by a set of procedural rules and conditions, which do not relate necessarily to inputs and outputs of each task, or to the nature of transformations performed within each task (especially for those performed by humans). Within a scientific workflow, conversely, data transformations and dependencies have a crucial importance. In other words, business workflows focus on control flows, described in terms of procedural rules and conditions; scientific workflows focus on data flows, described in terms of data dependencies [1, 48].

3. THE NESTED WORKFLOW MODEL

We propose a nested workflow model aimed at providing a fine-grained description of a visual analytics solution at different levels of abstraction and composition, and at any phase of design or validation. The main components of our workflow model are *tasks*, *users*, and *data* (Figure 1). They correspond to the domain blocks (also known as situations) in the NBGM by Meyer et al. [25], as well as to the vertices of the design triangle for visual analytics by Miksch and Aigner [26]. This workflow model allows designers to capture the structure of relationships and interdependencies between blocks, such as synchronization between tasks that are performed by the same technique or algorithm, or fusion of data sources and collaboration between users to solve a given problem. At the same time, the nested structure of each block enables the representation of identified or designed properties of that block at the different levels. In particular, *task blocks* have up to four levels of nesting, from the problem domain (the outermost level), trough the abstraction and the encoding, to the algorithms (the innermost level). Data blocks have up to two nesting levels: the domain problem (outer level), and the abstraction level (inner level). We name these blocks more properly *data/knowledge blocks*, as explained in Section 3.3. *User blocks* obviously represent situations at the problem domain level, but they are characterized in terms of prior knowledge, which can be abstracted into inner levels. It is worth noting that, when applying the workflow model to represent real scenarios, not all blocks must necessarily present all levels. For example, part of the workflow might consist of simple tasks, such as data preparation and data quality checks, processed by a generic user without involvement of domain knowledge.

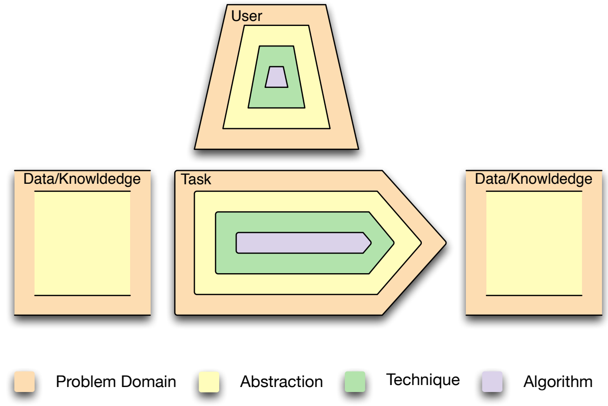


Figure 1: The three main components of the nested workflow model: tasks, data/knowledge blocks (input/output of the tasks), and users (who perform tasks, indipendently or in collaboration). The workflow components present up to four nesting levels [27]: problem domain, data and operation abstractions, visual encoding and interaction techniques, and algorithms.

By comparing our workflow model with other workflows, we observe that *tasks* are a common component of both business and scientific workflows, *users* correspond to the resource component of business workflows, while *data* represent the main focus of scientific workflows. Indeed, our workflow model for visual analytics is different from both business and scientific workflows in terms of interaction with participants and flow paradigm. In particular, it differs from business workflows because it focuses on data flow (and also extends it, see Section 3.3); it differs from scientific workflows because it considers only humans as actors.

3.1 Characterizing the problem domain: work practices and processes

The need of studying real settings (i.e., a given problem domain with a number of users, real datasets, and realistic tasks) has been identified as one of the challenges for validation of visualization [30]. Munzner's nested model [27], by eliciting its four levels, has also revealed that these levels have not received the same amount of attention by researchers: the problem domain level, in particular, is the most neglected one (with notable exceptions, both early [16, 39] and more recent ones [43, 45]). In the context of visual analytics research, Endert et al. [8] also highlight the need of recognizing analysts' work processes, and seamlessly fitting analytics into that existing interactive process. Lam et al. [21], considering seven scenarios for empirical studies in information visualization, note that only few works have addressed the Understanding Environments and Work Practices (UWP) scenario; therefore, they observe that studies about people and their processes are rare. Isenberg et al. [17], in their review on evaluation of visualization based on the seven-scenario classification, also note that very few papers have addressed the UWP scenario. They group the UWP scenario together with the Visual Data Analysis and Reasoning (VDAR) scenario, which is considered in an even lesser number of papers. UWP and VDAR scenarios are grouped becasue they both aim at understanding domain experts' analysis processes and practices. However, UWP focuses on understanding the current practices and is assimilable to a requirements analysis, while VDAR focuses on the use of a newly introduced visualization tool by a group of domain experts. In other words, UWP is traditionally seen as a scenario for formative evaluation, while VDAR is understood as a scenario for summative evaluation, according to the categories introduced by Ellis and Dix [7]. In our experience within applied research projects, we observed that the difference between UWP and VDAR scenarios has been getting smaller. Visualization and visual analytics have become more mature disciplines, commercial products exist and organization have started to adopt and use them. Domain experts do not provide only problems and data to designers and ask for solutions from the scratch; they might also show their existing solutions and ask for redesign, extension, or improvement. As a result, in our nested workflow model the distinction between identified and designed blocks is less strict, and also abstractions, techniques, and algorithms can be identified by observing users' analytical workflows already in place. Moreover, also in a pure UWP scenario (i.e., before the introduction of a running prototype or system), the inner levels of workflow components in our model can be populated with existing artifacts and interactions. Tory and Staub-French [39], for example, also characterize a problem domain in terms of artifacts (e.g., sticks, notes, and sketches), and manual interactions (e.g., five-finger pointing gesture to indicate an area, two-finger pointing gesture to indicate the distance between two points). In summary, the proposed nested workflow model aims at accompanying the NBGM, by complementing design guidelines with a mechanism which allows designers to understand the structure of existing workflows, identify components that can be improved (by introducing newly designed blocks or redesigning existing ones), and validating interventions by before/after comparisons.

3.2 Tasks

The central components of our nested workflow model are tasks. In the visualization literature, there is not even a consensus about the terminology, and researchers use several terms with overlapping or conflicting meanings: task, question, problem, objective, activity, action, operation. Rind et al. [32] try to untangle the confusion in terminology, by introducing a conceptual space of tasks along three dimensions: *abstraction* (domain specific/domain independent), *composition* (low level/high level), and *perspective* (how/why). Brehmer and Munzner [4] consider only abstract tasks, and identify three perspectives: *why*, *how*, and *what*. *How*, in particular, refers to the methods (visual encodings and interaction techniques) by which a task is executed, while *what* refers to its inputs and outputs. Our nested workflow model draws on the why/how/what characterization, but extends it also along the composition and the abstraction dimensions. In particular, the *why* refers to the operation abstraction level but also to the outer level of the problem domain; the nesting enables mapping between the two levels. The *how* is represented by the two inmost levels: visual encodings and interaction techniques, and algorithms. The *what* is represented by the *data/knowledge* blocks as the task inputs and outputs. The workflow model enables complex (de-)compositions of tasks beyond the simple task sequence, by exploiting several workflow patterns. Figure 1 shows a branching pattern (it might be a selection, or a parallelism, for example) and a merging pattern (a synchronization, for example). It is worth noting that the figure is intended as an conceptual illustration of the nested workflow model, while standard visual tools can be used to represent complex patterns, such as UML diagrams [41] or BPMN [28].

3.3 Data and Knowledge

Similarly to scientific workflows, the focus on data is important for visual analytics workflows. Data, therefore, represent the obvious input and output of visual analytics tasks. Input raw data can be complemented with explicit knowledge (metadata, ontologies), which can be exploited to assist the visualization [6], but also to drive the automated analysis and help users interpreting results [3, 9]. As noted by Chen et al. [6], "a computer representation of a piece of information or knowledge is just a particular form of data".

Tacit knowledge can also constitute both an input and an output of a visual analytics task. The value of visualization, indeed, lies in its capability to enable the acquisition of new knowledge based on the data, the visual perception, and the user's prior knowledge of the user [11, 42]. Similarly, during the visual analytics process, analysts formulate hypotheses on the basis of their prior knowledge about the problem domain, and gain new knowledge by verifying those hypotheses [34]. Tacit knowledge, and especially its increments, are obviously difficult to measure, observe, and represent. Specific methods of Cognitive Task Analysis (CTA) might be required to elicit tacit knowledge, such as interviews (i.e., asking people questions), self-reports (i.e., people talk about their behavior and strategies), observations of performance or task behavior, and automated collection of behavioral data [2]. If not externalized during the workflow, tacit knowledge remains specific to each user and, therefore, in our model it is represent within the user block (see Section 3.4). Nevertheless, in many circumstances tacit knowledge is externalized as part of the analytical process. Shrinivasan and van Wijk [35] note the importance of knowledge externalization within the analytical reasoning process, for example by annotating insights, hypotheses, and evidences. Wang et al. [44] describe knowledge conversions (internalization and externalization) within a visual analytics process.

Our nested workflow model, therefore, features two-level knowledge/data blocks, in order to represent not only input and output data, but also explicit knowledge that is available at the beginning of a task, or that is made available as an outcome of the task completion.

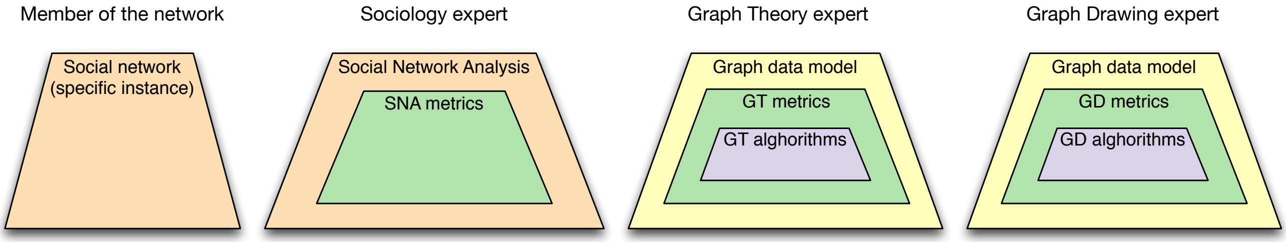


Figure 2: Example of user blocks in the domain of social networks, characterized at different levels.

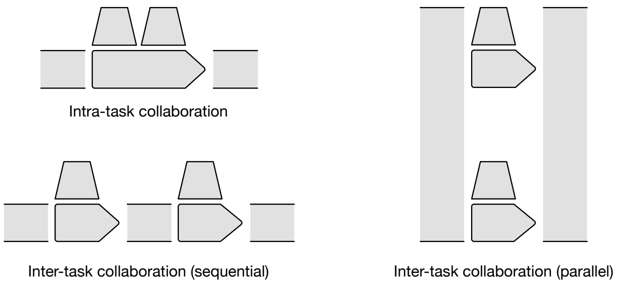
3.4 Users and collaboration

Users are an obvious component of the problem domain, and their understanding by designers is essential for the human-centred design process of interactive systems in general [14, 37], and visulization in particular [20]. Differently from scientific workflows, visual analytics systems make certainly use of automated methods for data analysis (e.g., data mining, machine learning), but the human "is always the ultimate authority in directing the analysis" [19, page 2]. The most recent directions in visual analytics, suggesting a shift from the human-in-the-loop to the human-is-the-loop paradigm [8], make the role of the user even more central. Sometimes, when systems are designed for a single user (or a single user group), design studies do not necessarily make an explicit reference to the user, since the user is identified with the problem domain itself. Conversely, when the design needs to address diverse users, designers need to characterize each of them, as well as their relationships with the other situations of the problem domain. In our model, users are characterized in terms of their tacit prior knowledge. In the context of visualization, by drawing on concepts from artificial intelligence, we distinguish two types of prior knowledge: operational knowledge (the knowledge of how to operate the system), and domain knowledge (the knowledge of how to interpret the content) [5]. These two types of knowledge span across the four nested levels. Let us illustrate it with an example from evaluation of visual analytics of social networks [36], see Figure 2. At the problem domain level, users can have some prior knowledge to interpret a social network, for example because they are sociology experts, or because the social network represent themselves and their friends. At the data and operation abstraction level, experts in graph theory have the knowledge to understand the modelling of data as a graph, and to compute graph-theoretic metrics such as the centrality degree or the network diameter. Nevertheless, only sociologist have the domain knowledge to interpret social network analysis metrics. Graph drawing experts have the operational knowledge at the visual encoding and interaction level, and therefore are able to switch to the most effective layout algorithm to solve a certain task (e.g., count the number of clusters). They might also have the operational knowledge at the algorithmic level to fine tune the layout parameters.

Therefore, in our model users are characterized in terms of their prior knowledge and are represented by four-level nested blocks, containing references to knowledge, data, operations, techniques, and algorithms. By using the workflow model, designers can quickly check whether users' tacit knowledge, complemented with available explicit knowledge at the task input, fulfils the knowledge requirements to accomplish a task, and can adjust possible mismatches.

Collaboration between users is a fundamental principle of visual analytics. According to the first recommendation by Thomas and Cook, designers should "build upon theoretical foundations of reasoning, sense-making, cognition, and perception to create visually enabled tools to support *collaborative* analytic reasoning about complex and dynamic problems" [38, page 6]. Golovchinsky et al. [10] provide a classification of collaboration in information seeking, an activity that presents analogies with the lower steps of the foraging loop in the sensemaking process as described by Pirolli and Card [29]. Collaborative information seeking can be described in terms of forms of collaboration, and roles in collaboration. The forms of collaboration include the level of mediation by the system, the concurrency (synchronous/asynchronous), and the location (shared space/individual space). The roles in collaboration create different configurations: collaboration between peers, between two experts of two different domains, between an expert and a novice of the same domain, between a domain expert and a search expert, and between a prospector (user interested in an overview) and a miner (user interested in details). Isenberg et al. [15] define collaborative visualization, classify it according to the space-time matrix [18] (co-located or distributed, synchronous or asynchronous), and delineate research directions. Ribarsky et al. [31] remark the importance of collaboration for analytical reasoning, and also note that researchers should identify which artifacts should be shared, in what form, and at what stage of the reasoning process. They list three types of collaboration: peers with different tasks; analysts with other stakeholders; analysts at different levels within the organization hierarchy. Heer and Agrawala [12] discuss design considerations for collaborative visual analytics, and describe the artifacts to be shared in terms of *deixis* (reference). The reference can be general, definite (e.g., by unique id), detailed (e.g., by attribute), or deictic (e.g., by pointing directly to something). Mahyar et al. [23] highlight the importance of externalization (for example, by taking notes) in collaborative analysis, and the lack of support in many visualization tools.

However, collaboration has not attracted much attention in visualization evaluation studies, with few exceptions (such as [24]). Lam et al. [21] identify two scenarios related to collaboration between users: the Communication Through Visualization (CTV) scenario, and the Collaborative Data Analysis (CDA) scenario. Isenberg at al. [17] in their survey found no studies for the CTV scenario, and very few studies for the CDA scenario.



Intra-task collaboration

Inter-task collaboration (sequential)

Inter-task collaboration (parallel)

Figure 3: The nested workflow model can represent different types of collaboration between two users, such as intra-task collaboration, inter-task collaboration in sequence, inter-task collaboration in parallel.

Our model enables the characterization of different aspects of collaboration in visual analytics workflows. First of all, it supports the characterization of multiple users, with different prior knowledge at all levels and different roles. Second, the inmost levels of task blocks can represent shared artifacts (interactive visualization, as well as non-computer-mediated artifacts). Last, it can model the assignment of tasks to different users, exclusively or in collaboration. In particular, since the task block is the central component of the nested workflow model, we introduce a task-based characterization of collaboration (Figure 3). We distinguish between intra-task collaboration (multiple users collaborate to complete the same task) and inter-task collaboration (each user completes one task, and data/knowledge are shared across the task boundaries). Different types of inter-task collaboration are possible, according to different task composition patterns (e.g., sequence, parallelism, iteration).

4. APPLYING THE MODEL

In this section, we demonstrate the application of the nested workflow model to two ongoing visual analytics research projects, discussing how its expressiveness helped us to understand users' environments, work practices, and visual data analysis reasoning.

4.1 Visual analytics of unexploded ordnance risks

In the context of an applied research project, we collaborate with a company whose work consists of identifying, assessing, and reporting unexploded ordnance risks from World War II, by collecting and analysing data from different sources, such as historical aerial imagery, current orthophotos, digital terrain models, militray records, and reports from historians and direct witnesses. In order to characterize the problem domain and understand analytics processes in place, we conducted semi-structured interviews [46] and contextual inquiries [13], and modelled our observation by using the nested workflow model.

Figure 4 shows an overview of the entire workflow (top), and a close-up of one of its parts (bottom). For clarity of exposition, also in the close-up view the model has been simplified and purged of unnecessary details.

We identified eight high-level tasks: retrieve images, retrieve events, filter images (twice), identify suspicious objects, ortho-rectify and geo-reference images, assess risks, and report findings. We observed that each high-level task can be decomposed into low-level tasks, showing different workflow patterns, in particular selection, iteration, and parallelism. Then we abstracted domain specific low-level tasks into domain-independent data operations.

We also observed a number of artifacts used to perform the tasks, such as geographic information system (GIS) with their interaction techniques, text and image databases with their information retrieval techniques and algorithms, electronic spreadsheets, simple image viewers, text editors, as well as emails, paper printouts, and hand-written annotations. In the close-up of Figure 4 (bottom), the technique level is missing. The reason is that in this part of the workflow only simple tools are used (such as text and image viewers), without any particular visualization, interaction, or automated analysis technique.

We identified three kinds of users: senior analysts, junior analysts, and image retrieval experts. We characterized them in terms of domain and operational knowledge. Both senior analysis and junior analysts have the knowledge to operate a GIS software and exploit its visual encodings and interaction techniques (e.g., semi-transparent layering of different images of the same area, or swiping of upper layers to show layers beneath, as well as procedures to ortho-rectify and geo-reference aerial images). We observed that senior analysts are capable of fast geographic orientation (e.g., finding the area of interest with precise, seamless panning and zooming interactions). Moreover, senior analysts have the domain knowledge to understand jargon of military reports and to interpret terrain features, e.g. to identify craters created by exploded bombs and correlate the crater diameter with the possible bomb types. Junior analysts have enough familiarity with aerial imagery to assess its quality, according to factors such as resolution, lighting conditions, or cloud cover. Image retrieval experts are capable of operating the company proprietary tools and algorithms.

By analyzing the workflow, we also identified different patterns of inter-task collaboration, such as sequence, parallelism, and also iteration (i.e., when an analyst cannot decide about an object identification, s/he asks colleagues to repeat the task and cross-check findings). We observed different types of knowledge externalization and sharing, such as annotations, collection of evidences for the final report, email exchanges and file transfers, and occasionally also deictic reference.

Overall, the nested workflow model allowed us to describe the existing analytic processes (in terms of tasks, users, and data/knowledge, at different nested levels) and to start the redesign of parts of the workflow (for example, by introducing interactive visualization techniques to speed up and optimize the information retrieval process, or by integrating computer vision methods, interactively supervised by the user, in order to improve identification and geo-referencing of suspicious objects). Morever, the nested workflow model provides a framework to track redesign interventions, and to plan adequate post-hoc comparisons and validations.

4.2 Visual analytics for medical shared decision making

This research project is focused on shared decision-making in the context of medical treatment of patients with atrial fibrillation. The management of this condition by anticoag-

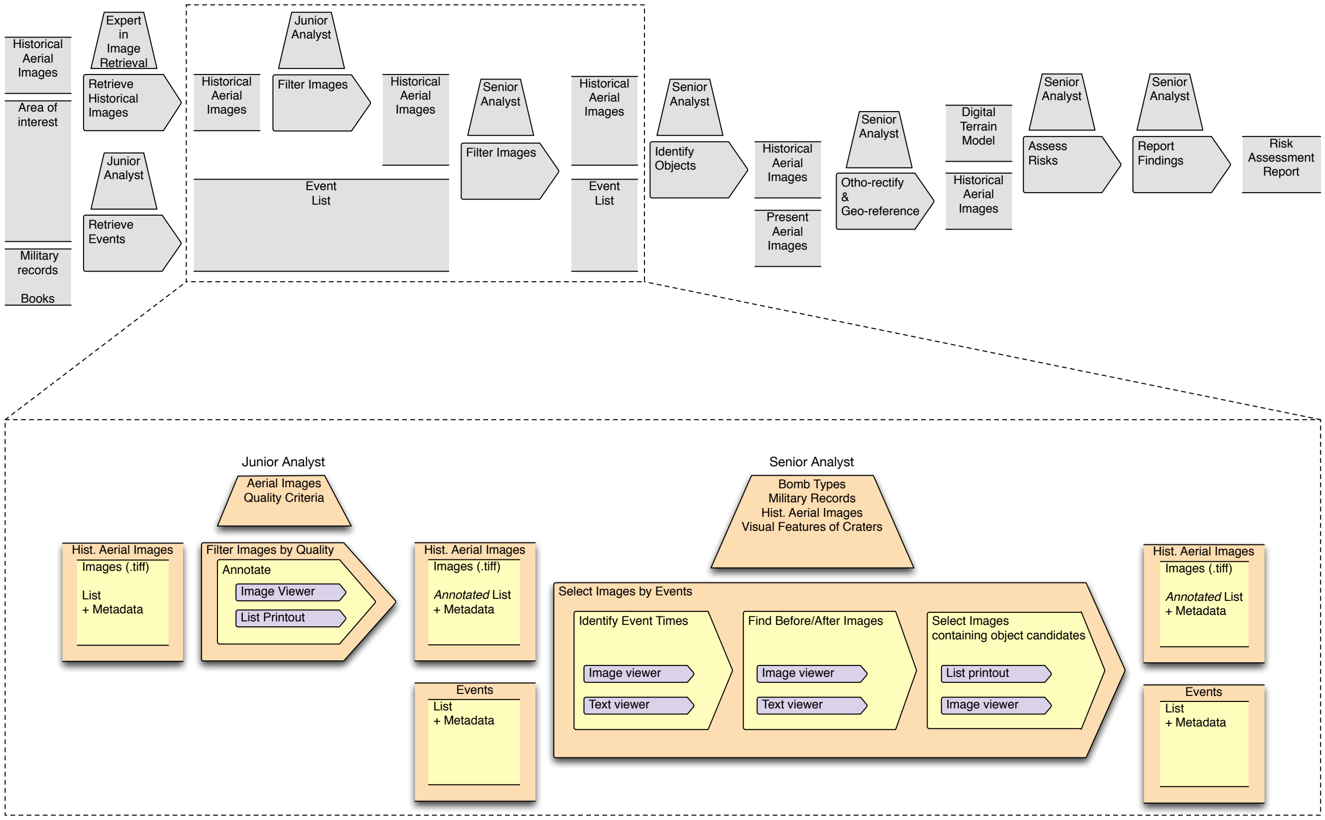


Figure 4: The nested workflow model applied to a study for visual analytics of unexploded ordnance risks. Top: overview; bottom: close-up of the *Filter* tasks.

ulant therapy is aimed at reducing the risk of strokes and other ischemic events, but at the same time it increases the risk of bleeding. Therefore, physicians need to inform their patients about the different treatment options and their implications and help them taking critical decisions. The aim of our project is to provide decision support, in the context of shared decision making, by means of a visual analytics system based on patient-level microsimulations of Markov models and graphical representation of simulated life paths [33]. We interviewed domain experts in the medical field as well as in medical decision-support systems, and applied the nested workflow model. The expressiveness and the flexibility of the model were particularly useful in this project, which has peculiar features.

The users are the cardiologist and the patient. The former is a domain expert, and has the operational knowledge to manage statistical models and the domain knowledge to understand medical concepts and understand patient condition data. The latter lacks the domain knowledge, but is the data owner and also the principal stakeholder. The large asymmetry in the prior knowledge makes also difficult to identify the scenario, which lays between Communication Trough Visualization (CTV) and Collaborative Data Analysis (CDA). However, the nested workflow model allowed us to capture the data and knowledge flows.

The main data in this circumstances are the transitions probabilities of the Markov models, representing the medical evidence derived from previous studies, but the visual encoding of this data is never shown to the patient. They are used to run simulations, whose results are then visually encoded and shown to the patient. Moreover, the doctor needs to elicit patients preferences in order to quality-adjusted life years (QALYs, a measure of life expectancy including both quantity and quality of lived years).

The nested workflow model helped us with the design and validation, by allowing us to represent different processes of visual communication, knowledge externalization, and shared analysis, facilitated by analytical methods and interactive visualizations.

5. DISCUSSION AND CONCLUSION

In this paper, we have introduced the nested workflow model. Our model builds upon Munzner's nested model [27], and inherits its benefits; in particular, it allows designers to identify the different levels and to consider specific threats to validity for each level. While Munzner's work can be seen as a theoretical and methodological framework, our contribution provides an operationalization of the nested methodology focused on the workflow perspective. By comparing the nested workflow model with the NBGM [25], we observe that the major commonality is the identification of blocks. In addition to NBGM data and task situations, our model also considers user situations; it can represent multiple users as well as different kinds of collaboration between them. Moreover, while in the NBGM the relationships between blocks are guidelines, representing design alternatives and recommendations, our model concentrates on the structure of the workflow in terms of tasks, execution patterns, responsible users (alone or in collaboration), data/knowledge flows and interfaces. Our preliminary experience of applying the model showed us that the redesign of a block (e.g., changing visual encodings, or integrating automated data analysis) sometimes leads to changes also in task execution order and user assignment; therefore, a representation of the visual analytics workflow also supports designers with validating the redesign and to compare before/after performances.

More in general, our workflow model is obviously intended to complement design and evaluation methods discussed in the visual analytics literature, and not to replace them. We are also aware that some aspects of our model might reflect common practices of visual analytics designers, but to the best of our knowledge they have not been documented as such. Therefore, we present our nested workflow model as a contribution towards a common framework for the community, to ease experience sharing and critical reflection. Future work comprises specifying a standard graphical representation of the conceptual model, as well as eliciting standard methods to summarize and compare nested workflow instances.

Acknowledgments

The authors wish to thank Luftbilddatenbank Dr. Carls GmbH, Stefania Rubrichi, and Silvana Quaglini.

This work was partially supported by the Bridge 1 Programme of the Austrian Research Promotion Agency (FFG), grant number 850695, project DeVisOR.

6. REFERENCES

[1] A. Barker and J. van Hemert. Scientific workflow: A survey and research directions. In R. Wyrzykowski, J. Dongarra, K. Karczewski, and J. Wasniewski, editors, *Parallel Processing and Applied Mathematics: 7th International Conference, PPAM 2007, Gdansk, Poland, September 9-12, 2007 Revised Selected Papers*, pages 746–753, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[2] R. R. H. Beth Crandall, Gary Klein. *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. A Bradford Book. The MIT Press, 1 edition, 2006.

[3] 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, pages 12:1–12:8, New York, NY, USA, 2013. ACM.

[4] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, Dec 2013.

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

[6] M. Chen, D. Ebert, H. Hagen, R. S. 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.

[7] G. Ellis and A. Dix. An explorative analysis of user evaluation studies in information visualisation. In *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '06, pages 1–7, New York, NY, USA, 2006. ACM.

[8] A. Endert, M. S. Hossain, N. Ramakrishnan, C. North, P. Fiaux, and C. Andrews. The human is the loop: new directions for visual analytics. *Journal of Intelligent Information Systems*, 43(3):411–435, 2014.

[9] 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, editors, *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association, 2015.

[10] G. Golovchinsky, P. Qvarfordt, and J. Pickens. Collaborative information seeking. *Computer*, 42(3):47–51, Mar. 2009.

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

[12] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. *Information Visualization*, 7(1):49–62, 2008.

[13] K. Holtzblatt and S. Jones. Contextual inquiry: A participatory technique for system design. In D. Schuler and A. Namioka, editors, *Participatory design: Principles and practices*, pages 177–210. Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1993.

[14] International Organization for Standardization. Ergonomics of human-system interaction – Part 201: Human-centred design for interactive systems. Standard, ISO, Geneva, CH, Mar. 2010.

[15] P. Isenberg, N. Elmqvist, J. Scholtz, D. Cernea, K.-L. Ma, and H. Hagen. Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, 10(4):310–326, 2011.

[16] P. Isenberg, A. Tang, and S. Carpendale. An exploratory study of visual information analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1217–1226, New York, NY, USA, 2008. ACM.

[17] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, Dec 2013.

[18] R. Johansen. *GroupWare: Computer Support for Business Teams*. The Free Press, New York, NY, USA, 1988.

[19] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann, editors. *Mastering the Information Age. Solving problems with Visual Analytics*. The Eurographics Association, 2010.

[20] O. Kulyk, R. Kosara, J. Urquiza, and I. Wassink. Human-centered aspects. In A. Kerren, A. Ebert, and J. Meyer, editors, *Human-Centered Visualization Environments: GI-Dagstuhl Research Seminar, Dagstuhl Castle, Germany, March 5-8, 2006, Revised*

Lectures, pages 13–75. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[21] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, Sept 2012.

[22] J. Liu, E. Pacitti, P. Valduriez, and M. Mattoso. A survey of data-intensive scientific workflow management. *Journal of Grid Computing*, 13(4):457–493, 2015.

[23] N. Mahyar, A. Sarvghad, and M. Tory. Note-taking in co-located collaborative visual analytics: Analysis of an observational study. *Information Visualization*, 11(3):190–204, 2012.

[24] N. Mahyar and M. Tory. Supporting communication and coordination in collaborative sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1633–1642, Dec 2014.

[25] M. Meyer, M. Sedlmair, P. S. Quinan, and T. Munzner. The nested blocks and guidelines model. *Information Visualization*, 14(3):234–249, 2015.

[26] S. Miksch and W. Aigner. A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286 – 290, 2014.

[27] T. Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, Nov 2009.

[28] Object Management Group. Business Process Model and Notation BPMN Version 2.0. Normative document, OMG, Needham, MA, Jan. 2011.

[29] 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*, volume 5, pages 2–4, 2005.

[30] C. Plaisant. The challenge of information visualization evaluation. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '04, pages 109–116, New York, NY, USA, 2004. ACM.

[31] W. Ribarsky, B. Fisher, and W. M. Pottenger. Science of analytical reasoning. *Information Visualization*, 8(4):254–262, 2009.

[32] A. Rind, W. Aigner, M. Wagner, S. Miksch, and T. Lammarsch. Task cube: A three-dimensional conceptual space of user tasks in visualization design and evaluation. *Information Visualization*, 2015.

[33] 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.

[34] 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.

[35] Y. B. Shrinivasan and J. J. van Wijk. Supporting the analytical reasoning process in information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 1237–1246, New York, NY, USA, 2008. ACM.

[36] M. Smuc, P. Federico, F. Windhager, W. Aigner, L. Zenk, and S. Miksch. How do you connect moving dots? Insights from user studies on dynamic network visualizations. In W. Huang, editor, *Handbook of Human Centric Visualization*, pages 623–650. Springer New York, New York, NY, 2014.

[37] D. Stone, C. Jarrett, M. Woodroffe, and S. Minocha. *User Interface Design and Evaluation (Interactive Technologies)*. Morgan Kaufmann, 2005.

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

[39] M. Tory and S. Staub-French. Qualitative analysis of visualization: A building design field study. In *Proceedings of the 2008 Workshop on BEyond Time and Errors: Novel evaLuation Methods for Information Visualization*, BELIV '08, pages 7:1–7:8, New York, NY, USA, 2008. ACM.

[40] W. van der Aalst, A. ter Hofstede, B. Kiepuszewski, and A. Barros. Workflow patterns. *Distributed and Parallel Databases*, 14(1):5–51, 2003.

[41] W. M. P. van der Aalst and K. M. van Hee. *Workflow management: models, methods, and systems*. Cooperative information systems. MIT Press, 1st edition, 2002.

[42] J. J. van Wijk. The value of visualization. In *VIS 05. IEEE Visualization, 2005.*, pages 79–86, Oct 2005.

[43] M. Wagner, W. Aigner, A. Rind, H. Dornhackl, K. Kadletz, R. Luh, and P. Tavolato. Problem characterization and abstraction for visual analytics in behavior-based malware pattern analysis. In *Proceedings of the Eleventh Workshop on Visualization for Cyber Security*, VizSec '14, pages 9–16, New York, NY, USA, 2014. ACM.

[44] 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, 2009.

[45] K. M. Winters, D. Lach, and J. B. Cushing. Considerations for characterizing domain problems. In *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, BELIV '14, pages 16–22, New York, NY, USA, 2014. ACM.

[46] L. E. Wood. Semi-structured interviewing for user-centered design. *interactions*, 4(2):48–61, Mar. 1997.

[47] Workflow Management Coalition. Terminology & Glossary. Specification, WFMC, Winchester, UK, Feb. 1999.

[48] U. Yildiz, A. Guabtni, and A. H. H. Ngu. Business versus scientific workflows: A comparative study. In *2009 Congress on Services - I*, pages 340–343, July 2009.