

A Two-stage Framework for Designing Visual Analytics System in Organizational Environments

Xiaoyu Wang

Wenwen Dou

Thomas Butkiewicz

Eric A. Bier

William Ribarsky

UNC Charlotte

UNC Charlotte

University of New Hampshire

Palo Alto Research Center

UNC Charlotte

ABSTRACT

A perennially interesting research topic in the field of visual analytics is how to effectively develop systems that support organizational users' decision-making and reasoning processes. The problem is, however, most domain analytical practices generally vary from organization to organization. This leads to diverse designs of visual analytics systems in incorporating domain analytical processes, making it difficult to generalize the success from one domain to another. Exacerbating this problem is the dearth of general models of analytical workflows available to enable such timely and effective designs.

To alleviate these problems, we present a two-stage framework for informing the design of a visual analytics system. This design framework builds upon and extends current practices pertaining to analytical workflow and focuses, in particular, on incorporating both general domain analysis processes as well as individual's analytical activities. We illustrate both stages and their design components through examples, and hope this framework will be useful for designing future visual analytics systems. We validate the soundness of our framework with two visual analytics systems, namely Entity Workspace [8] and PatViz [37].

KEYWORDS: Design Theory, Visual Analytics, Human Computer Interaction.

INDEX TERMS: H5.2 [[Information Interfaces And Presentation (e.g., HCI)]: User Interfaces (D.2.2, H.1.2, I.3.6) Theory

1 INTRODUCTION

Organizational visual analytics systems represent an emerging technology, shown to be useful in facilitating domain analytical reasoning and insight discoveries through interactive visual interfaces [13, 58]. As the Visual Analytics (VA) field strides forward, research has been applied to various domains, leading to the development of diverse tailored systems. As summarized by Thomas and Kielman [35], many existing systems are useful in helping users tackle challenging domain analytical processes, such as finance [13] and infrastructure management [58]. However, a perennially recurring research topic is the identification of a general design framework to facilitate the development of visual analytics systems.

Much like the progression in other empirical sciences (e.g. HCI, InfoVis), many initial models [14, 41, 46] suggest that the establishment of a general design framework is significant. The objective for such a VA design framework is threefold: firstly, the framework must inform designers to systematically incorporate the support for domain analytical processes. Secondly, the framework should provide a basis for designers to evaluate their system and further help them identify a cohesive technology transition progress, from system design and implementation to its release and deployment [49]. Finally, the framework must serve an educational purpose, contributing to the identification of potential course materials that are necessary to educate others regarding the field of VA [1].

However, constructing a convincing and appropriate design framework is challenging. The framework must be validated against existing systems and more importantly, it must give

researchers and designers new ideas regarding how to evaluate and improve their own work.

Given the need to incorporate successes from diverse VA systems, it is difficult to generate a framework that can summarize and instruct all the design requirements from a top-down perspective. High-level VA design frameworks like [14, 41, 44, 55] are certainly of great value. Nonetheless, little specific guidance or recommendation is currently available to articulate the boundaries within which particular design assumptions apply, leaving the system design to be solely based on designers' prior experience. For example, how does a designer know which analysis method is suitable to characterize an organization? Are there components that a designer should follow to systematically incorporate a domain analytical process? Further, what recommendations exist that specify the appropriate methods for supporting both general and individual analytical workflows? Encouraged by the discussions in the VAC consortium 2010 [1], the authors of this paper resolved to document their experiences with VA framework design.

In the pursuit of a general VA design framework, we grounded the conceptualization of such a framework in our collaborations with three organizations [58, 59, 64], and materialized the framework by integrating guidelines from existing literature [41], [61] with our experience. We started by designing the framework to fit the emerging characteristics of the field of visual analytics, and further enriched its components based on our design practices and the novel ideas presented by others. To preserve the generalizability of the framework, we studied a broad selection of literature on modeling organizational analytical processes and designing information systems from research fields such as Organizational Learning (OL), Knowledge Management (KM), Machine Learning (ML), Human Computer Interaction (HCI), and Information Visualization (InfoVis). Although much of our work can be applied to these domains, in this paper, we focus on presenting and validating a two-stage framework that is primarily specific to the VA field.

We hope this framework will be useful in designing and evaluating VA systems. We illustrate both stages and their design components through examples. We further use external evidence [40] to evaluate the usefulness and novelty of the framework. In some cases, we compare components of our framework with existing work. In other cases, we present recommendations that are largely unique to the design of VA framework. Due to space constraints, we cannot review all visual analytics systems nor exemplify all the components of such systems.

In the following section, we discuss the two design fundamentals of a VA framework that designers must consider. Existing models and frameworks related to these fundamentals are further described and discussed. In the subsequent sections, we present the details of our two-stage design framework and its related design components. We organize these components into two design stages. First, we discuss the Observation and Designing stage (Section 4.2), which designers should follow to characterize the targeted domain and support its general analytical processes. While the first stage is valid in presenting the synthesis of the majority of domain analysis activities, individual differences aren't captured in this stage. Consequently, we describe the need for a "feedback" process to integrate the

individual's analytical practices into VA systems. In particular, we present the second, User-centric Refinement stage (Section 4.3), in which designers find recommendations to customize the visual analytics system to support individual ways in performing analysis. We design both stages to be complementary and hope they can collectively provide recommendations that are informative for developing an organizational VA system.

2 RELATED WORK

Similar to the emphasis in the information system design [19], many researchers have recognized the importance for VA designs to conflate task activities within both the organizational level (high-level holistic task workflow) and the individual level (specific tasks or operations) [44, 46]. The pioneers of visual analytics---researchers from data analysis, InfoVis, and HCI---contributed invaluable application designs, algorithmic thinking, and engineering traditions. Quite early on, the intelligence community stressed the application of analysis models and theories of sense-making processes when designing VA systems [11]. As shown in the sense-making loop, Card and Pirolli [45] presented a theoretical basis for understanding and portraying the analytical discourse that an analyst performs. Later on, cognitive scientists showed interests in this emerging field and introduced a scientific design approach, focusing on the value of empirical observations of analysis processes and users' performances [27].

Recently, influences from ML, HCI and KM contributed to the establishment of the methodological grounds for VA research. These influences have broadened the scope for VA designs from focusing on creating visual interfaces, to accommodating broader and more complex analytical needs between individuals, teams and organizations. The contemporary visual analytics field is multidisciplinary and is rapidly maturing thanks to the collective contributions from these diversified research areas.

2.1 The Fundamentals of a VA Design Framework

Like many empirical sciences (e.g. HCI and InfoVis), the field of visual analytics does not solely research on existing technologies, styles of interaction, or interface solutions. The design of organizational VA systems extends beyond merely the final visual interfaces and representations. While these interface features are undeniably essential considerations in visual analytics system development, they are not the major concern of this work.

As shown in Figure 1, the core foundation of a VA framework emphasizes the *integration* of general domain analytics processes, the *visual facilitation* of domain analytical tasks and finally, the *customization* for individual's analytical workflows. Essentially, the design of a VA system is *analysis-centric*, in that it needs to encapsulate the organizational analysis discourse [41] and support its related reasoning tasks and user behaviors [44].

In particular, such design needs to consider and alleviate potential incompatibilities and challenges that could affect users' acceptances of, and reactions to, an interactive VA system. This design also needs to incorporate supports for various domain analysis tasks, properly match the nature of the task and the support from the system, provide logical organization of targeted data, utilize accurate statistics to meaningfully transform data, and most importantly, guarantee consistency between analytical workflows and VA system operations.

Human-centered design is another significant factor in VA design. Recent research and practices in HCI have shown that users' affective reactions and their holistic experiences with technology are becoming increasingly important [2, 61]. As suggested by many empirical studies, a better understanding of various human cognitive, affective, and behavioral factors in the

context of problem solving processes, are essential in designing a more successful analytical system [61].

The central tenet of human-centered design is to fully engage domain users during the modelling and design process of a VA system, ensuring their requirements and demands are clearly understood and conveyed to designers. Such design is a bi-directional process: elicitation of system requirement demands VA designers to communicate fluently in the same "analysis language" of the targeted domain; it also requires designers to introduce the merits of visual analytics to influence the organization and help improving its existing analysis processes.

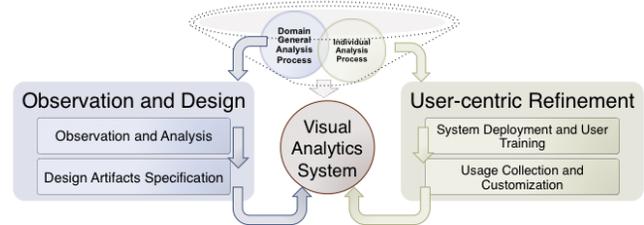


Figure 1: This illustrates the general design flow in our proposed design framework, including support for both general analysis workflow (Left) and individual analytical processes (right).

3 MOTIVATION

We consider both the generalization of core domain analysis and human-centered design as two fundamentals of a VA design framework. Such framework, therefore, needs to comply with these fundamentals in that: 1) it must reveal the generalizability of visual analytics in encapsulating and facilitating domain analysis processes, and more importantly 2) the framework must clearly instruct a systematical development process that guarantees the efficacy and validity of a customizable visual analytics system.

Previous design models [15, 55] emphasize using data to tell the story. In the data drive visualization design, researchers focus on accommodating the nature of the data. They emphasize utilization of mathematical and statistical methods in deducing the information embedded in a dataset. One of the earliest theorists in InfoVis, Jacques Bertin had noted that the understanding of deduction of relationships is a matter of permutation [6]. Bertin further proposed a synoptic that differentiated between ordered, reorderable, and topographic data, established retinal qualities of marks that would allow viewers to differentiate between marks, and provided recommendations for representing data as arrays of marks, histograms, and curves based on its dimensionality.

In addition, Chen et al. [14] in their position paper on knowledge-assisted visual analytics system proposed a high-level design pipeline. This pipeline focuses on utilizing visualization to help users transfer data from the computational space to information and knowledge in the perceptual and cognitive space. They suggest the need for VA infrastructures to support the development of visualization and to transfer such data to information and knowledge, helping further our understanding and enhancing visualization technology. While this pipeline provided a clear conceptual design direction for VA systems, it has been too high-level to be informative for actual development.

Last but not least, Munzner [41] proposed a nested model that focused on the use of validation in guiding the visualization designers through the system design processes. Such nested model presented a unified approach to visualization design and illustrated effective methods in evaluating the designed systems. While it is not directly targeted at addressing challenges in VA, this model has influenced the development of this paper.

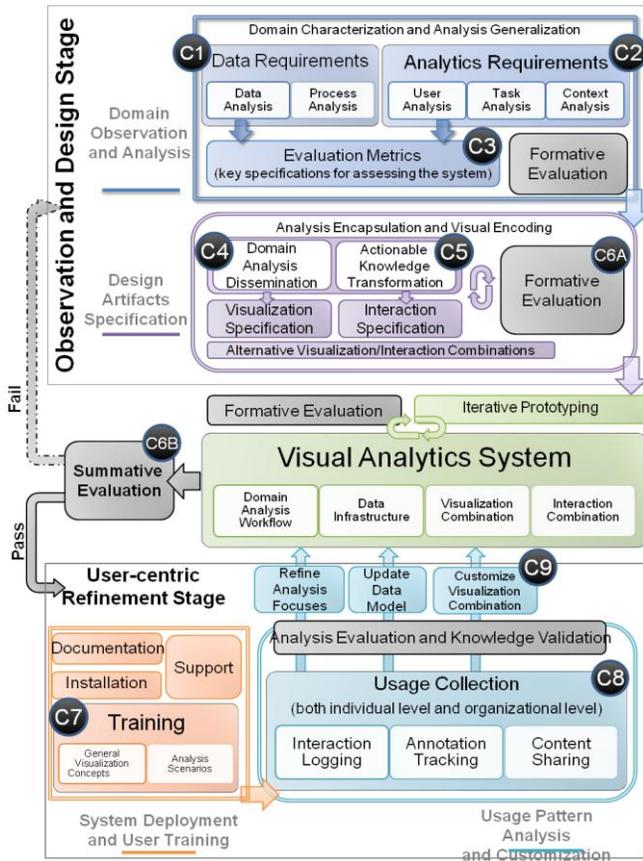


Figure 2: The overview for our proposed framework, including both the Observation and Designing stage and the User-Centric Refinement stage. Dashed line is potential development directions, not necessarily required for every development. C_x (1 to 9) refers to individual design component, where x being the component ID.

4 A TWO-STAGE FRAMEWORK

Following Zimmerman's [62] definition of design research as "an intention to produce knowledge and not the work to more immediately inform the development of a product", instead of illustrating a specific implementation practice, we propose a two-stage framework for systematically designing VA systems in organizational environments. Similar to Klein's [31] view on developing an information system, our framework is conceptualized based on the incorporation of both general analysis workflow and individual analytical processes (see Figure 1); and it is further materialized to conflate the two above fundamentals in a cohesive manner to augment the design of a VA system.

In this section, we present the details of our two-stage VA design framework (see Figure 2). First, we introduce the methodology used in deducting this framework. Then, we illustrate both design stages and their components through examples. We further validate the necessities and usefulness for these components by comparing them with research in other relevant fields (e.g. HCI and InfoVis).

4.1 Methodology

The design process of this framework is intertwined with actual implementation practices. This process emphasizes bootstrapping both theoretical design aspects and practical experiences to construct a coherent descriptive design framework.

The implementation practices present essential functional components to be incorporated into the framework. They further

provide a testing ground to verify and validate the design framework. Our design framework is grounded in actual system development with three groups of professionals in different organizational settings: bridge-asset managers in The U.S. Department of Transportation [58], who propose and execute strategic bridge maintenance plans; business analysts from Xerox [59], who retrieve and analyze documents for information essential to the operation of the business; and network-asset managers from Microsoft, who monitor and maintain the operations of back-end server farms [64]. The developments of these VA systems were carried out through close examination of domain users' analytic workflows and interviews of their analytical actions that are taken to achieve each task.

Our framework is constructed along with these collaborations to instruct and enrich our implementation practices. It details the natural progression of designing a visual analytics system and presents it in a cohesive manner. We started by extending current models (e.g., the nested model [41]) and frameworks pertaining to analytical workflow and focused, in particular, on investigating its effects on the design of visual analytics systems for organizational environments. As reported in [56], we then iteratively categorized our design experiences from these collaborations into a more general organizational analysis workflow, and used it to inform future system designs. In doing so, we identified and enriched the design components of our framework by comparing the commonality and differences between individual analytical domains. We further validated and encapsulated them into a coherent two-stage design framework, as illustrated in Figure 2.

4.2 Stage I: Observation and Characterization

The first stage in our framework addresses the question of incorporating the general domain analytical processes. Designers must identify the essential analytical processes in the targeted domain, and disseminate and transform these processes into tangible visualization and interaction specifications. In the conceptual level, this stage is similar to the nested design model; however, our framework is more detailed on the actions taken in each design step. The construction of this stage has also been greatly influenced by Holtzblatt and Beyers' work on Contextual Design [66]. Particularly inspired by their proposed guidelines on participatory design, we emphasize the importance of data-and -analytics-requirement elicitations during this first design stage.

In the following sections, we introduce design components that can inform designers with the methods necessary to characterize the domain and incorporate general domain analytical processes in their system designs.

4.2.1 Domain Observation and Analysis

Objective: *Characterize the target domain's analytical processes and specify evaluation metrics for validating designs.*

Following the recommendations from Munzner [41] and Ribarsky et al. [46], we consider the *characterization* of an organization's analysis activities as the most crucial step in the design of visual analytics systems. At this step, designers must establish common vocabulary, and interact with domain users to acquire data and analytical requirements for the targeted domain. Designers must also emphasize the derivation of evaluation metrics in assessing the efficacy of the designs. We echo the call from Scholtz [47] that such metrics must be derived based on clear elicitation of both the data and analytical requirements.

As detailed below, we present two main components (analytics requirements and evaluation metrics) that are unique to the characterization of the domain analytical discourse. For eliciting data requirements ($C1$), designers can refer to previous work by van Wijk [55] and Bertin [6].

4.2.1.1 Component (C2): Analytical Requirements

This component describes how to map problems and users' needs from a specific domain into more generic analytics requirements, including analysis environments, high-level task flows, and social aspects within the organization. In particular, analytics requirements can be generated based on context analysis, user analysis, and task analysis [61].

Context Analysis: As noted in organizational design and HCI research [9, 36, 61], context analysis is widely used and is considered to be the initial step when approaching organizational users. Specifically, in the context of VA design, we consider context analysis in a more focused scope, eliciting the organization settings (e.g. technical, analysis and collaboration settings), in which the systems will be used. It provides a constraint to ensure important factors that may affect the usability of a product are considered, and provide insights to develop a strategic plan for the system design. In addition, designers should use context analysis to examine whether and how users interact within or across organizations would impact or even change the design of a visual analytics system. To perform a thorough context analysis, we followed Zhang et al. [61] and recommended the use of the following three specific analyses:

Technical Context Analysis focuses on identifying the technical specifications within an organization. It answers questions such as: What is the preferred display, hardware, or OS in the organization? What are users' technical affordances in accepting advanced visual representation (e.g. Parallel Coordinates)? Designers must use this analysis to establish technical baselines and shape their design strategies accordingly. More importantly, they should follow the results of this analysis to match designs with user's technical skills, minimizing any potential cognitive overheads. For example, while both Taste [59] and Document Card [53] were designed to perform similar document activity analysis, given the constraints on the user's hardware and their technical needs, Taste was designed for a more confined screen space, with both detail view and aggregated overviews.

Operational Context Analysis is useful in mapping out the relationship between resources (data, techniques, etc.) and restrictions. It answers questions like: What entities and resources are used in task operations? Where tasks are carried out? What operational policies are used in this organization? For example, the researchers, who developed WireVis [13], elaborated their data integration and acquisition processes specific to, and enforced by, the financial institution's policies.

Organizational Context Analysis should be used to identify the information pertinent to an individual's analytical needs (e.g. personnel structure and reporting hierarchy), and more importantly, to portray the dynamic analysis flows that are essential to the organization's operations [10, 28]. This analysis typically addresses questions such as: Which streams of data need fuse together? Where to retrieve information for certain analytical tasks? How to share individuals' analysis with others, and finally, in what way could one collaborate with others to reach business decisions? Such cases can be seen in the design research by Convertino et al. [16], in which the authors elaborated their practices of identifying and capturing general analytical workflows for two organizations.

User Analysis: User analysis is used to ensure the information and characterizations are accurate and explicit to the target users. This analysis provides the designers with perspectives on the different categories of domain users who will ultimately use the visual analytics systems. Conceptually, user analysis may seem obvious; in practice however, it is not trivial. As suggested by Dillon et al. [20], user analyses are typically highly context

sensitive and vary from one organization to another. Thus, to perform a proper user analysis, designers must actively engage domain users and elicit the design requirements through extensive interactions with these users. Based on previous research in OL [24] and HCI [61], we summarize three typical focuses of the user analysis (Table 1).

Table 1. The information provided by user-analysis

Demographics Data	Occupation, Organizational position, Specific task focuses, Computer skill, and Experience with similar analytical systems.
Task Related Factors	Job characteristics, Frequency of analytical tools used for the tasks, and Usage constraints and preferences.
Personal Traits	Cognitive styles, Affective traits, Work styles, and Skill sets or capabilities

Task Analysis: Once user and context analysis have passed, task analysis is conducted to specify the tasks and workflows in an organizational environment. In general terms, task analysis focuses on analyzing and articulating the nature of analytics tasks that are normally performed in a given organization [24].

Specific to the context of VA design, we consider task analysis in a more focused scope, stressing the specification of actionable knowledge [5] (detailed in Section 4.2.2) that is directly associated with, and utilized in, the domain analytical workflows. The goal of this analysis is for designers to identify the needed fine-grain analytical actions (what to do and how to perform the task), the task structure (how one task leads to another), and the strategy of task organization (which tasks are concurrent).

We recommend designers to consider approaching task analysis at two levels: individual and organizational. Within the individual level, designers should start by identifying the analytical goals meaningful to an individual's work within the operational. They should decompose these goals into task activities or actions that users must do to interact with the system. At the organizational level, designers need to sequence individual tasks into coherent high-level analytical workflows. They should further specify cross-process tasks that are commonly applicable to multiple workflows and identify actions that are unique to certain workflows. The system should be designed to directly support the individual tasks within the organization workflows as well as their high level contexts.

Following ethnographic methodologies [65], typical methods to perform the task analysis are semi-structured interview [59], surveys [58], and on-site observations.

4.2.1.2 Component (C3): Evaluation Metrics

The evaluation metric specifies the expected analysis goals from the domain users, and presents key features that a VA system should incorporate and be evaluated by. The specific measures or quantitative aspects of the metrics are typically determined based on the aforementioned analyses, formative evaluation, as well as the goals and constraints of the organization.

Following the evaluation methods commonly used in the intelligence community [30], we recommend the use of qualitative evaluation metrics for evaluating the analytical utility of VA system (Table 2). These evaluation metrics correspond to the five concerns of VA systems [47] and provide designers with detailed measures to assess how the visualizations facilitate analysis, how users interact with the visualizations, and what supports the analysis environment. Designers should refer to the evaluation metrics (C3) typically used as benchmarks for both formative evaluation (e.g. designers' assessment on prototypes) [22] (C6A) and summative evaluation (e.g. end-users' assessment on the system) [50] (C6B). Note that, given the domain differences, not all measures are required at the same time. Designers must choose

the most appropriate metrics based on the domain specifications. Examples of evaluation metrics of use can be found in [48].

Table 2. A list of areas of VA evaluation concern

Concerns	Description	Sample Measures
Situation Awareness	VA system supports the analysts' knowledge on performing domain specific analytical tasks	Ability to track the changes of information; Provide contextual analysis environments; Self-descriptiveness of actions
Collaboration	VA system enables communication and information sharing between collaborators	Ability to share evidence; Support intuitive communication; Capable to reveal information flows
Interaction	VA system provides sufficient visualization and interaction combinations to facilitate domain analytical processes	Suitability for the task; Controllability; Support customization
Creativity	VA system supports the flexible and diversified analytical processes for individual analysts	Support individual tasks; Effective in searching for analytical results; Ability to show high quality of analysis solutions
Utility (Analytical Process)	VA system fits in analysts cognitive strength and reduces the cognitive workload on analysts	Easy to use; Engaging; Comply with exiting technical context; Conformity with user expectation

4.2.2 Design Artifacts Specification

Objective: Disseminate the high-level domain tasks to tangible design artifacts and transform them into system functionalities.

The task activities and general workflows identified in the previous design step are useful in characterizing a particular domain; however, they are often too high-level to provide any specific recommendations for actual system designs. Therefore, designers must follow this step to disseminate the domain analysis processes into tangible design artifacts, and transform them into visualization and interaction specifications.

In the following sections, we present two major components that can advise designers on what to look for when searching for tangible design artifacts, and help them identify the best methods for incorporating these artifacts into system designs. To provide the designers a concrete understanding of both the dissemination and transformation process, we further illustrate key action knowledge and its matching design considerations based on our design experience with the aforementioned three organizations. This process is detailed separately in [56].

4.2.2.1 Component (C4): Analysis Dissemination

The first action in this component is to search for tangible artifacts that could help breakdown the high-level semantic tasks of a specific domain. Designers should derive tangible artifacts based on two requirements: (1) they need to be concrete enough for practical VA system designs, and more importantly (2) such artifacts must be consumable by users without introducing considerable cognitive overhead.

Enlightened by the Theory of Action [5], we recognize the usefulness of describing target artifacts as *Actionable Knowledge*. Actionable knowledge is explicit symbolic knowledge that is well accepted in organizational environments. It is typically presented in the form of tradeoffs-for-action or rules [39], which are used to instruct domain user's actions when performing a task. It presents a pragmatic view of knowledge utilization and application towards specific analytical ends [12]. Actionable knowledge

details the relationships between domain analytical tasks and their related key actions, such as what entity to examine, or which person to communicate with in the process of collaborations [8]. The nature of actionable knowledge fits well with our two requirements in that: (1) it represents the fine-grained elements of each analytical task, and thus is quite instructive for the design of a VA system; (2) it is extracted from domain users' knowledge actions, and therefore can be consumed without additional cognitive overhead.

As exemplified in [5, 12], there are many approaches to model actionable knowledge. Given the VA designers' advantage of a close working relationship with actual domain users, designers should adopt the domain-driven modeling process and ground their search for actionable knowledge in interviews and surveys. The designers should encourage interviewees to envision the hypothetical process of carrying out their analytical tasks with their familiar tools, in their familiar working environments. They should also encourage interviewees to think about additional functions that might be useful, but not yet available in any of the tools they typically use. Specifically, designers need to ask about the fine-grained task actions that are used in the users' daily practices, and understand what the essential tools are and how these tools are utilized in each task action. In doing so, designers can identify key actionable knowledge that is informative and detailed enough to incorporate specific domain analysis processes.

4.2.2.2 Comp. (C5): Actionable Knowledge Transformation

To determine the likelihood of domain users accepting a visual analytics system's functionalities, designers must elicit the specifications for visual interfaces and interactions needed after the identification of the above actionable knowledge. Both formative evaluations and iterative prototyping serve an invaluable role in this process. These processes essentially help designers encapsulating domain users' actionable knowledge into system functions, and inform them of the insights into the critical functionalities required to implement the specific VA system.

Given the diverse design context in different organizations, designers need to carefully select prototyping methods. For example, given the requirement of a deployable product within 10 weeks, we applied the evolutionary prototyping [43] method in our collaboration with Xerox [59]. This prototyping method guarantees more design iterations and, more importantly, allows the deployment of a robust system in a structured manner. In contrast, when engaged in a long-term collaboration, we used a functionality-based prototyping process to design our bridge management system for USDOT [58].

In addition, designers should rely on formative evaluations to identify defects in function designs and refine their design decisions. In fact, we agree with Munzner [41] and recommend the iterative use of formative evaluations throughout the entire visual analytics system design and development process. We encourage the designers only form final design decisions based on their experiences gained from these evaluations.

4.3 Stage II: User-Centric Refinement

The second stage in this framework focuses on incorporating the domain users' individual analytical processes. We echo the call of Kindlmann [36] and Silva et al. [53] that, in order to recreate and extend specific visualization results, knowing the complete process of how the results are generated is just as important as the techniques used and the outcome. The process of recording how a user interacts with a visualization is sometimes referred to as provenance tracking, which is defined by Anderson et al. [4] as

“the logging of information about how data came into being and how it was processed.”

To succeed in the system refinement stage, designers must first perform summative evaluation and deploy the designed system to domain users. They then need to collect and analyze actual usage data from domain users and provide methods to customize the VA systems to encapsulate individual’s analytical process.

Recent work by Scholtz et al. [49] has presented an informative pipeline for deploying visual analytics systems to domain users. Their work covered a broad range of components in the deployment process, including the system deployment step in our framework. Thus, we are not restating this process (C7) here, and recommend designers to seek relevant information in their paper.

In the following sections, we focus on discussing details for the Usage Pattern Analysis and Customization process, which is key in capturing and incorporating individual’s analytical processes.

4.3.1 Usage Pattern Analysis and Customization

Objective: Support individual task routines and analysis preferences, enable individuals to collect analytical findings, and establish organizational collaboration.

In keeping with the need to customize a visual analytics system, designers must enhance their design to collect users’ usages of the system and incorporate individuals’ analytical processes. Particularly, there are two components (usage collection and system customization) that are necessary in this design process.

4.3.1.1 Component (C8): Usage Collection

In this component, we describe two useful methods designers can use to collect usage data in a VA system. These two methods include an implicit approach (logging users’ interactions), and an explicit approach (tracking users’ annotations). Depending on the organizational and operational context, designers should choose the most appropriate method to collect usage data. This data could then be used for analyzing user’s analysis behaviors (e.g. data focuses, view focuses, etc.), and preparing statistics for the customization process detailed in the following sections.

It is worth noting that rather than focusing on concluding a definitive set of usage collection technologies, we emphasize explaining the utility of this component in contributing to the overall design framework.

Annotation Tracking and Content Sharing: Annotation tracking is yet another method to capture the shared explicit knowledge between different users in an organization environment [11, 42]. We refer *annotation* to the process in which users externalize their findings (e.g. data correlation, outliers, patterns or trends) within the visualizations. Compared to interaction logging that focuses on implicitly capturing users’ analysis processes, annotations place the users in the analysis center by explicitly tracking and sharing their findings. Through annotation, domain users can properly attach semantic meanings to their analysis findings, and further analyze, evaluate, reuse, and exchange these findings for their collaborative decision-making.

HCI [29] and KM [42] research suggests that the key to successful annotation tracking is the symbolic content that can be used to inform users’ analysis process, share their knowledge, and build consensus of decision-making through the use of a computer system. Based on previous research, we categorize the design of an annotation tracking mechanism into two levels: the sharing of static annotations level (e.g. snapshots, bookmarks, or comments), and the exchanging of dynamic annotations level (e.g. parameters of a visualization [32]). The details of these two mechanisms are summarized and compared in Table 3.

It is worth noting that this explicit-collecting method requires users to manually attach semantic meanings to analytical findings

[26]. Designers should be aware of the potential introduction of interruptibility during the users’ analysis process. We are currently researching on more automated annotation methods (i.e. semi-defined annotations templates) to reduce such interruptions [57].

Interaction Logging: Interaction is increasingly seen as central in representing individuals’ analysis processes within VA systems [44]. We use the term “interaction” in the broad sense defined by Yi et al. [63]: “the dialogue between the user and the system as the user explores the data set to uncover insights.” It is through the interactive manipulation of a visual interface that knowledge is constructed, tested, refined, and shared [44]. In turn, a considerable amount of information regarding a user’s analysis process with a visualization tool is captured by interactions. Indeed, Dou et al. [21] discovered that high-level semantic reasoning processes can be recovered by asking novices to merely analyze experts’ interactions logs. With similar goals in mind, other researchers have proposed capturing user interactions on different levels, ranging from low-level mouse events (Glassbox [18]) to high-level task-specific actions (Gotz et al. [26]).

Based on previous research and our experiences in designing interaction-logging mechanisms [57, 59], we categorize this logging mechanism into three levels (as presented in Table 4). The benefit of such categorization is that designers can select appropriate logging elements based on their own design goals.

Table 3. The comparison list between sharing static annotation and exchanging dynamic annotations

Sharing Mechanism	Content	Efficiency	Effectiveness	Information Sharing Flow
Share Static Annotation	Static Image; Textual Information; Drawing;	Easy to construct; Can add on to existing VA systems	More effective in a small-to-mid-size group	Typically one-way. Info. Comes from original analysts to shared with other colleagues
Exchange Dynamic Annotation	Parameters that can be applied to another instance of the VA sys	Needs to modify the existing VA system.	Support larger collaboration teams and departments	Bi-directional: both original analysts and their peers can collectively modify and extend the analysis results

Table 4. A list of categorized interaction logging methods

Log Focus	Log Elements	Examples
Tracing Details of Analysis Sessions	Low level event (e.g. mouse click, key stroke)	Jeong et al. [33]
Replay Key Analysis Frames	Visual State (e.g. visualization parameters)	Jankun-Kelly et al. [32] Shrinivasan et al. [52]
Reconstruct User’s Analysis Processes	Task-level actions and Contextual information.	Gotz et al. [26] Dou et al. [21]

4.3.1.2 Component: Usage Analysis

In our framework, both interaction logging and annotation tracking serve the same goal, which is to provide inputs for VA systems to analyze users’ analysis behaviors. Many researchers have demonstrated designs using these methods to collect usage data. For example, by encoding users’ behaviors in a visual analytics system, Gotz et al. [25] presented a visual analytics system that can make appropriate suggestions to the user by analyzing the captured analytical processes. In addition, through the use of pre-defined scripting language, the Czsaw system [34] has shown capabilities in capturing and reapplying users’ analysis processes to similar tasks.

Regarding associating interaction logs with high-level reasoning process, Dou et al. [21] validated such a possibility in the specific context of financial analysis. Nonetheless, few guidelines were available on establishing the correlation between usage collection and the users' analysis processes, nor to recommend methods to analyze the collected data. Expanding on these findings, we further categorize the existing automatic usage analysis methods into three groups, namely general statistical analysis, graph theory, and machine learning.

4.3.1.3 Component (C9): System Customization and Knowledge Validation

From an organization's perspective, the design of a customizable VA system must address the following three needs: firstly, the designed system should have the ability to rearrange sequences and combinations of analytical components to support individual analyses. Secondly, the system should allow individuals to collect analytical findings and trace their analysis trails which led to these findings. Finally, such system needs to establish a collaborative environment that enables users to communicate their analytical findings. As shown in Table 5, we propose three methods to individualize the system, including refine analysis focuses, update data model, and customize visualization combination.

Table 5. Three system customization methods.

Method	Description & Examples
Update Data Model	Based on user's data focus, modify and update the underlying data model. E.g. the VA needs to prioritize the more frequently used statistics based on usage histories.
Customize Visualization Combinations	Rearranging the visualization combination based on the users' interaction logs and annotation histories. Built upon a modular design, the VA system should adjust the primary and entry views based on usage histories.
Refine Analysis Focuses	Utilizing the recorded annotation, the VA system needs to record the important analysis focuses for a user. It needs to guide the users toward their analysis goals through interactive guidance.

Analysis Evaluation and Knowledge Validation: Due to individual experiences and understanding, different experts have their own ways of performing analysis processes. Their views of an analytical process may be imprecise, duplicated, and even conflicted with the organization's generic analytical workflow. Therefore, it is a concern that, if new analysis processes or knowledge is not carefully validated, customizing the VA system with unrelated or incorrect knowledge could potentially degrade the value of such system.

While currently there is no existing literature specifically focused on this validation process, we propose investigation of the *costs* involved with customization. Echoing the evaluation research [47], such costs of customization could be a combined factor of cost of interaction (Lam [38]), cost of visualization (suggested by Amar et al. [3]), and cost of cognitive overload (proposed by Green et al. [27]). The designed visual analytics system needs to apply thresholds to control the cost, and maximize the cost/benefit value to determining whether or not to incorporate the new knowledge.

Therefore, while designers should follow these methods to customize their designs, we emphasize the importance of validating *costs* of these customizations before incorporating them into the system. It is necessary for designers to consider these costs before allowing users to customize their system. A two-step process can be helpful: The first step is to quantify the cost of customization and transform it into a tangible threshold for each visualization customization request. Any updates under that threshold will be automatically updated, while customizations that beyond that threshold will have to pass onto the second step:

where the users manually accept or dispute the candidates for system customization. Several existing systems have begun to explore such validation process. For example, the KEF framework [17], presents the user with the ability to review the automatically suggested update, and allows them to accept or dispute it based on user's preferences; also, the HARVEST system [51] actively tracks the visual changes in a system, and allows users to revisit or revert the visual representations interactively.

5 EXAMPLES

To provide concrete examples on how our framework could be implemented, we analyze two existing visual analytics systems in terms of our framework. We validate the soundness of our framework with two well-received systems, namely Entity Workspace [8] and PatViz [37]. To explicitly compare their design decisions with our framework, we indicate the same process by placing a mark *C_x* in Figure 2 (x referring to the component number shown in each stage).

5.1 Entity Workspace

Entity Workspace was first introduced in 2006 [8]. As an interactive visual analytics system, it combines user interface and entity extraction technologies to build up an explicit model of important entities (people, places, organizations, etc.) and their relationships. This system helps analysts find and re-find facts rapidly and gain awareness of connections between entities in a highly interactive environment. This VA system was designed to incorporate the support for general analysis, as proposed in our first design stage. Through a longitudinal collaboration with intelligence analysts, the designers elicited general analytical processes (e.g., re-find important facts and "connect the dots" between different entities) by performing analyses on both data (*C1*) and analytics requirements (*C2*). They further disseminated (*C4*) and transformed (*C5*) these high-level tasks into specific design features (e.g. snap-together relationship and a graph model for gathered intelligence), and iteratively implemented these features with feedback from analysts. These design activities complete the first stage in our framework.

At the second stage, the implemented system was evaluated (*C6*) with analysts and then deployed to analysts for long-term assessment (*C7*). Acting on feedback from analysts, the designers extended the functions of Entity Workspace to support collaborative intelligence analysis [7]. They provided five advanced collaborative features to support information sharing, particularly sharing of the immediate products of users' analysis (e.g. annotations, evidence, and document collections) (*C9*). The designers further implemented a recommendation mechanism (*C8*) to customize the analysts' workspaces. Through encapsulating the user's preferences into a graph model (*C8*), Entity Workspace can automatically suggest to analysts unforeseen entities of interest.

In summary, we consider Entity Workspace valuable external evidence in exemplifying the efficacy of our design framework. Eight of the nine design components from our framework are utilized in the design of the Entity Workspace. In particular, its annotation and recommendation features illustrate the usefulness of the proposed User-Centric Refinement stage (Stage II).

5.2 PatViz

PatViz [37] was designed to bridge the gap between retrieving and analyzing patent information by providing seamless integration of derived insights in consecutive query refinement. It aims to address the long existing scalability issues in the analysis of patent

information. PatViz was built as a graphical front-end for a set of different patent search engines, and provided a visual environment allowing interactive reintegration of insights into subsequent search iterations.

The design of PatViz came from requirements formalized by interviews with patent specialists from the PATExpert consortium. During the implementation of PatViz, the designers first performed detailed *user analysis* and *context analysis* to gain general understanding of the patent analysis domain. These analyses also helped explicitly explain the roles of patent analysis related professionals and their existing workflow and tools. The designers' extensive analysis of the challenging data scalability issues in patent search falls into the elicitation of the data requirements (C1).

Through interactions with patent analysts, the designers depicted the general patent analysis loop (C2) and its related three analytical processes, namely patent retrieval, patent result set analysis, and patent detail analysis. They further broke down these analytical processes into the essential system design components (C4), resulting in a querying system, a multitude of visual result set representations, and the linkage between them (C5).

While the design of PatViz heavily focused on the first stage in our framework, it also covered components in the User-Centric Refinement stage (Stage II). The researchers conducted two informal field studies to evaluate the scalability and accuracy of PatViz (C6), and utilized users' feedback to enrich the design of the system. Specifically, PatViz utilizes the Boolean integration language to implicitly record and reapply users' search histories (C8), and also to explicitly enable users to store and share search statements. PatViz could further analyze this usage information and customize users' visualization combinations (C9). The designers purposely implemented PatViz to involve the users in the loop of deciding which customized views to keep, demonstrating the utility of our proposed knowledge validation process (C9).

In summary, seven out of nine design components from our proposed framework were used in the design of the PatViz.

6 LIMITATIONS

There are limitations to this research that must be addressed. Generalizability of these design considerations is limited because this research was conducted within only three organizations. This paper attempted to mitigate local biases by increasing the number of participants. Nevertheless, different training backgrounds, personal preferences, and project time constraints could engender different analytical conditions.

Moreover, this research characterizes the domain analytical workflow through interviews and surveys, which generally are self-reported by participants. This research was also limited, in that it modeled the analytical workflow from a retrospective perspective, whereas Browns et al. [9] demonstrated that problem spaces and solutions are established and change dynamically in interactions with people and the environment. Therefore, the understanding of domain analysis and actionable knowledge is constrained to the users' general way of performing tasks.

Finally, while our framework has been evaluated by our previous formal studies on Taste [59] and informal case studies on IRSV [58] and OpsVis [64], this research is limited by its evaluations with domain experts. Developing evaluations, strategies, and methodologies to accurately assess the effectiveness of a visual analytics system is challenging. To this point, this paper still in search of a clear view of the best evaluation approach; the design of recommendations for

evaluating a visual analytic system is an intriguing direction for future research.

This paper recognizes these limitations and considers the support of organizational analysis processes as an important visual analytics research topic. The concluded design considerations illuminate the role that a visual analytics system plays in such complex problem-solving environments.

7 FUTURE WORK

This paper contributes to the establishment of a two-stage design framework for visual analytics. But some components in this framework still require further investigation. The uncovered complexity of this framework implies that there is a vast amount of work that must be done before a final framework is complete. In general there are three categories of future work: (1) expansion of interactive reasoning modeling capabilities for the framework; (2) understanding the cost of customization; and (3) establishing the evaluation foundation of visual analytics.

7.1 Expanding the interactive reasoning modeling capabilities

A first step in this direction is to expand the interactive reasoning modeling capabilities for our design framework. The analysis of the identified relevant domains of research could be deepened, and other relevant domains may be discovered. This expansion of the interactive modeling process could be used to refine and improve the incorporation of individual's analytical processes in a visual analytics system.

This research is interested in considering whether externalizing such domain knowledge and reapplying it into customized visualizations would be feasible for enhancing domain decision-making processes. Although there is no definitive way to achieve complete knowledge transfer, existing research has demonstrated how to incorporate visualization with domain specific knowledge [23]. To achieve similar knowledge externalization, a tight integration of the visualization with an ontological knowledge structure was proposed to interactively capture and store the user's interactions and translate them into domain knowledge [60]. This externalization could further be used in training new managers, communicating with others, and reporting decisions.

In addition, this work intends to investigate additional analysis methods for the automated analysis of user's interaction logs and annotations. For example, hidden Markov models could be used for data where segments are not explicitly defined but can be learned based on the original data sequence. These potential additions, combined with the general approach of blending automated and multi-view, interactive visual analysis, open the door to new insights that can help model the domain users' reasoning processes.

7.2 Understanding the costs in VA customization

The validation process of the cost for customization is of great importance in the "feedback" loop for the proposed design framework. As described in Section 4.3, this process emphasizes verification and validation of customization requests that are generated based on analyzing users' analytical behaviors. Key in this process is the identification of measures that can determine the costs of customizing a system. Currently, there is no existing research that addresses this recognized problem.

In the future, one of the most important research directions for this work is to continue investigating the measures for the costs of customization. This direction emphasizes the search for visual and/or interaction parameters that can be used to quantify such costs. Specifically, this research would focus on creating a

combined factor to attach costs to system customization. A first step in this direction could be to survey the existing literatures (e.g. visual analytics, InfoVis and HCI) for theoretical foundations for the cost of customization. The analysis of the identified relevant domains of research could be deepened, and other relevant domains may be discovered. This expansion of the theoretical foundation could be used to refine and improve the definition of the cost in customizing a VA system.

7.3 Establishing the evaluation foundation of VA

General evaluation recommendations for the assessment of the proposed visual analytics framework have not been solved here. Future work is needed to complete this research. The final solution may be some mixture of the utilization of internal evidence and external evidence [40]. Both present a coherent perspective to evaluate the framework, by placing it into the evolution of the visual analytics field. The evidence is collected to support speculation that such a mixed solution may be more useful than any one solution in isolation.

These general evaluation recommendations should not only focus on the assessment of the functionalities of a visual analytics system. One needs to verify the utility of a designed system, and validate how properly the implemented functions are in facilitating domain analysis process. These recommendations should also emphasize measuring the knowledge gain for the domain users.

8 CONCLUSION AND CONTRIBUTION

This paper has presented three years of iterative design efforts in searching for a visual analytics design framework. There are three primary contributions presented in this paper: first, we propose a two-stage framework for facilitating domain users' workflows through integrating their analytical models into interactive visual analytics systems. This design framework illustrates general design recommendations that, when followed, empower visual analytics systems by bringing the users closer to the center of their analytical processes. By integrating the analytical models into interactive visual analytics, users could directly interact with data in real time and make analytical decisions in a customized reasoning environment.

The second contribution of this paper is providing a general basis to bridge research and industry regarding design and development. It connects academic research on visual analytics to industrial organizations, and showcases the utility of organizational visual analytics systems. It not only provides industrial collaborators with concrete ideas about the impact that a VA system can provide, but also suggests a practical framework and considerations for designing visual analytics system.

Finally, this paper provides academia with a theoretical approach to understanding and designing visual analytics systems. It encourages researchers to search for, and establish a foundation for visual analytics design. This paper also serves an educational purpose and is intended to influence course syllabuses and materials for teaching visual analytics research.

We hope that by proposing our framework, we can promote a serious discussion of design considerations critical for producing effective knowledge-assisted visual analytics systems. We will continue to evaluate and refine our framework with current and future collaborators. In addition, we hope that this framework and its design guidelines will lead to potential impacts in today's organizational environments.

REFERENCES

- [1] VAC Consortium 2010, <http://www.vacommunity.org>.
- [2] R. Agarwal and E. Karahanna. Time flies when you are having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly*, 2000.
- [3] R. A. Amar and J.T. Stasko.. Knowledge Precepts for Design and Evaluation of Information Visualizations. *IEEE Transactions on Visualization and Computer Graphics* 11, 4 (July 2005), 432-442.
- [4] E. W. Anderson, C.T. Silva, J.P. Ahrens, K. Heitmann and S. Habib; "Provenance in Comparative Analysis: A Study in Cosmology,". *Computing in Science & Engineering*, vol. 10, no.3, pp.30-37, May-June 2008.
- [5] C. Argyris and D. A. Schon. *Theory in Practice: Increasing Professional Effectiveness*. Jossey-Bass, 1992.
- [6] J. Bertin. Graphics and graphic information processing. In *Readings in information visualization*, Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman (Eds.). Morgan Kaufmann Publishers Inc., San Francisco, CA.
- [7] E. Bier, S. Card, and J. Bodnar. Entity-based collaboration tools for intelligence analysis. In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 99–106, 2008.
- [8] E. A. Bier, E. W. Ishak, and E. Chi. Entity quick click: rapid text copying based on automatic entity extraction. In *CHI '06 extended abstracts on Human factors in computing systems*, pages 562–567, New York, USA, 2006. ACM.
- [9] J. S. Brown and P. Duguid. *Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovation*. *Organization Science*, 2(1):40–57, 1991.
- [10] T. Bucher, A. Gericke, and S. Sigg. Process-centric business intelligence. *Business Process Management Journal*, 15(408-429), 2009.
- [11] S. Card, J. J. Thomas, and K. A. Cook. Illuminating the Path: The Research and Development Agenda for Visual Analytics. Chapter 2. *National Visualization and Analytics Ctr*, 2005.
- [12] P.R. Carlile, A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science*, Vol. 13, No. 4, July-August 2002, pp. 442-455.
- [13] R. Chang, M. Ghoniem, R. Kosara, W. Ribarsky, J. Yang, E. Suma, C. Ziemkiewicz, D. Kern, and A. Sudjianto. Wirevis: Visualization of categorical, time-varying data from financial transactions. *VAST 2007. IEEE Symposium on*, pages 155–162, 30 2007-nov. 1 2007.
- [14] M. Chen, D. Ebert, H. Hagen, R. S. Laramée, R. van Liere, K.-L. Ma, W. Ribarsky, G. Scheuermann, and D. Silver. Data, information, and knowledge in visualization. *IEEE Comput. Graph. Appl.*, 29(1):12–19, 2009.
- [15] E. H. Chi, J. Pitkow, J. Mackinlay, P. Pirolli, R. Gossweiler, and S. K.Card. Visualizing the evolution of web ecologies. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 400–407. ACM Press/Addison-Wesley Publishing Co., 1998.
- [16] G. Convertino, S. Kairam, L. Hong, B. Suh, and E. H. Chi. Designing a cross-channel information management tool for workers in enterprise task forces. In *Proceedings of the International Conference on Advanced Visual Interfaces, AVI '10*, pages 103–110, New York, NY, USA, 2010. ACM.
- [17] A. J. Cowell, M. L. Gregory, E. J. Marshall, and L. McGrath. Knowledge encapsulation framework for collaborative social modeling. In *AAAI Spring Symposium on Technosocial Predictive Analytics*, 2009.
- [18] P. Cowley, L. Nowell, and J. Scholtz. "Glass box: An instrumented infrastructure for supporting human interaction with information". In *Proceedings of the 38th Annual Hawaii International Conference on System Sciences, 2005. HICSS '05*, pages 296c–296c, January 2005.
- [19] G. B. Davis and M. H. Olson. "Management information systems: conceptual foundations, structure, and development (2nd ed.)". McGraw-Hill, Inc., New York, NY, USA, 1985.9
- [20] A. Dillon and C. Watson., "User analysis in HCI—the historical lessons from individual differences research". *International Journal of Hum.-Comput. Stud.*, 45:619–637, December 1996.
- [21] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang. "Recovering reasoning processes from user interactions". *Computer Graphics and Applications, IEEE* , vol.29, no.3, pp.52-61, May-June 2009.
- [22] L. Fuchs and D. Fuchs. "Effects of systematic formative evaluation: a meta-analysis". *Exceptional Children*, v53 n3 p199-208 Nov 1986.

- [23] S. Garg, J. Nam, I. Ramakrishnan, and K. Mueller. "Model-driven visual analytics". In *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 19–26, Oct. 2008.
- [24] J. Gerlach and F.-Y. Kuo. "Understanding human computer interaction for information systems design". *MIS Quarterly*, 2008.
- [25] D. Gotz and Z. Wen. "Behavior-driven visualization recommendation". In *IUI '09: Proceedings of the 13th international conference on Intelligent user interfaces*, pages 315–324, New York, NY, USA, 2009. ACM.
- [26] D. Gotz and M. X. Zhou. "Characterizing users' visual analytic activity for insight provenance". *Information Visualization* 8, 1 (January 2009), 42–55.
- [27] T. Green, W. Ribarsky, and B. Fisher. "Visual analytics for complex concepts using a human cognition model," *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, vol., no., pp.91–98, 19–24 Oct. 2008.
- [28] J. Heer, N. Kong, and M. Agrawala. "Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations". In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1303–1312, New York, NY, USA, 2009. ACM.
- [29] J. Heer, J. D. Mackinlay, C. Stolte, and M. Agrawala. "Graphical histories for visualization: Supporting analysis, communication, and evaluation". *IEEE TVCG*, 14(6):1189 – 1196, 2008.
- [30] R. Heuer. "Psychology of Intelligence Analysis". Pherson Associates., 2007.
- [31] R. Hirschheim and H. K. Klein. "Four paradigms of information systems development". *Commun. ACM*, 32:1199–1216, October 1989.
- [32] T. Jankun-Kelly, K.-L. Ma, and M. Gertz. "A model and framework for visualization exploration". *IEEE TVCG*, 13(2):357–369, March/April 2007.
- [33] D. H. Jeong, W. Dou, H. Lipford, F. Stukes, R. Chang, and W. Ribarsky. "Evaluating the relationship between user interaction and financial visual analysis". *Visual Analytics Science and Technology, 2008. VAST '08. IEEE Symposium on*, pages 83–90, Oct. 2008.
- [34] N. Kadivar, V. Chen, E. Dunsmuir, D. abd Lee, C. Qian, J. Dill, C. Shaw, and R. Woodbury. "Capturing and supporting the analysis process". In *VAST 2009. IEEE Symposium on*, 2009.
- [35] D. A. Keim, F. Mansmann, and J. Thomas. "Visual analytics: How much visualization and how much analytics". *SigKDD Explorations Journal*, 2009.
- [36] G. Kindlmann. "Lack of reproducibility hinders visualization science". *IEEE Visualization Compendium*, IEEE CS Press, 2006.
- [37] S. Koch, H. Bosch, M. Giereth, and T. Ertl. "Iterative integration of visual insights during patent search and analysis". *IEEE Symposium on Visual Analytics Science and Technology*, 2009.
- [38] H. Lam. "A Framework of Interaction Costs in Information Visualization". *IEEE Transactions on Visualization and Computer Graphics* 14, 6 (November 2008), 1149–1156..
- [39] M. L. Markus, A. Majchrzak, and L. Gasser. "A design theory for systems that support emergent knowledge processes". *MIS Quarterly* 2002, 26(3):pp. 179–212,
- [40] D. D. McDonald,. "Internal and external evidence in the identification and semantic categorization of proper names". In *Corpus processing for lexical acquisition*, Branimir Boguraev and James Pustejovsky (Eds.). MIT Press, Cambridge, MA, USA 21–39..
- [41] T. Munzner. "A Nested Model for Visualization Design and Validation". *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (November 2009), 921–928..
- [42] I. Nonaka and H. Takeuchi. "The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation". Oxford University Press, USA, May 1995.
- [43] S. Overmyer. "Revolutionary vs. evolutionary rapid prototyping: balancing software productivity and HCI design concerns". In: *Proceedings of the 4th International Conference on Human-Computer Interaction*, 1991.
- [44] W. A. PIKE, J. Stasko, R. Chang, and T. A. O'Connell. "The science of interaction". *Information Visualization*, 8:263–274, 2009.
- [45] P. Pirolli and S. Card. "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis". *Proc. Int'l Conf. Intelligence Analysis*, 2005.
- [46] W. Ribarsky, B. Fisher, and W. M. Pottenger. "Science of analytical reasoning". *Information Visualization*, 8(4):254 – 262, 2009.
- [47] J. Scholtz. "Beyond usability: Evaluation aspects of visual analytic environment". *Proceedings of the IEEE Symposium On Visual Analytics Science And Technology (VAST)*, 2006.
- [48] J. Scholtz. "Developing qualitative metrics for visual analytic environments". *The Proceedings of BELIV '10*, 2010.
- [49] J. Scholtz, K. A. Cook, M. A. Whiting, D. Lemon, and H. Greenblatt. "Visual analytics technology transition progress". *Information Visualization*, 8:294–301, December 2009.
- [50] M. Scriven. "Evaluation thesaurus". Sage Publications., 1991.
- [51] Y. B. Shrinivasan and D. Gotz. "Connecting the dots with related notes". In *CHI '09: Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, pages 3649–3654, New York, NY, USA, 2009. ACM.
- [52] Y. B. Shrinivasan and J. J. van Wijk. "Supporting the analytical reasoning process in information visualization". In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1237–1246, New York, NY, , 2008. ACM.
- [53] C.T., Silva, J. Freire, S.P. Callahan; "Provenance for Visualizations: Reproducibility and Beyond," *Computing in Science & Engineering* , vol.9, no.5, pp.82–89, Sept.-Oct. 2007. doi: 10.1109/MCSE.2007.106
- [54] H. Strobel, D. Oelke, C. Rohrdantz, A. Stoffel, D. A. Keim, and O. Deussen. "Document cards: A top trumps visualization for documents". *Visualization and Computer Graphics, IEEE Transactions on* , vol.15, no.6, pp.1145–1152, Nov.-Dec. 2009
- [55] J. van Wijk. "The value of visualization". *Visualization*, 2005. VIS 05. IEEE, pages 79–86, Oct. 2005.
- [56] X. Wang, T. Butkiewicz, W. Dou, E. Bier and W. Ribarsky: "Designing knowledge-assisted visual analytics systems for organizational environments". Accepted in *VINCI 2011*, Hong Kong, China, 2011.
- [57] X. Wang, S.-E. Chen, E. Hauser, and W. Ribarsky. "iMonitor: Architecture of web-based collaborative visual analytics system for bridge management". In *Transportation Research Board 90th Meeting*, 2010, Washington DC. USA.
- [58] X. Wang, W. Dou, S.-E. Chen, W. Ribarsky, and R. Chang. "An interactive visual analytics system for bridge management". *Computer Graphics Forum*, 29:1033–1042, 2010.
- [59] X. Wang, B. Janssen, and E. Bier. "Finding business information by visualizing enterprise document activity". In *Proceedings of the International Conference on Advanced Visual Interfaces, AVI '10*, pages 41–48, New York, NY, USA, 2010. ACM.
- [60] 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 and Graphics*, 33(5):616 – 623, 2009.
- [61] P. Zhang, J. Carey, D. Te'eni, and M. Tremaine. "Integrating human computer interaction development into the systems development life cycle: a methodology". *Communications of the Association for Information Systems (Volume 15, 2005)*, 2005.
- [62] J. Zimmerman, J. Forlizzi, and S. Evenson. "Research through design as a method for interaction design research in HCI". In *Proceedings of the SIGCHI conference on Human factors in computing systems, CHI '07*, pages 493–502, New York, NY, USA, 2007. ACM.
- [63] Ji Soo Yi; Youn ah Kang; Stasko, J.T.; Jacko, J.A.; , "Toward a Deeper Understanding of the Role of Interaction in Information Visualization," *Visualization and Computer Graphics, IEEE Transactions on* , vol.13, no.6, pp.1224–1231, Nov.-Dec. 2007.
- [64] Fisher, D.; Maltz, D.A.; Greenberg, A.; Wang X.; Warncke, H.; Robertson, G.; Czerwinski, M.; , "Using visualization to support network and application management in a data center," *Internet Network Management Workshop, 2008. INM 2008. IEEE* , vol., no., pp.1–6, 19–19
- [65] Spradley, J.P. "The Ethnographic Interview", *Wadsworth Group/Thomas Learning*, 1979.
- [66] H. Beyer and K. Holtzblatt. "Contextual Design: Defining Customer-Centered Systems". Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.