Defining and Applying Knowledge Conversion Processes to a Visual Analytics System

Xiaoyu Wang, Dong Hyun Jeong, Wenwen Dou, Seok-won Lee, William Ribarsky and Remco Chang

Abstract

Knowledge-assisted visualization has been a fast growing field because it directly integrates and utilizes domain knowledge to produce effective data visualization. However, most existing knowledge-assisted visualization applications focus on integrating domain knowledge that is tailored only for specific analytical tasks. This reflects not only the different understandings of what "knowledge" is in visualization, but also the difficulties in generalizing and reapplying knowledge to new problems or domains. In this paper, we differentiate knowledge into two types, tacit and explicit, and suggest four conversion processes between them (internalization, externalization, collaboration, and combination) that could be included in knowledge-assisted visualizations. We demonstrate the applications of these four processes in a bridge visual analytical system for the US Department of Transportation and discuss their roles and utilities in real-life scenarios.

Key words: Knowledge-assisted visualization, Tacit Knowledge, Explicit Knowledge, Knowledge Conversion Processes, Visual Analytics

1. Introduction

The incorporation of knowledge into the process of solving analytical tasks is a fast emerging area in visualization. With the insights and reasoning artifacts that knowledge-assisted visualization provides, the experts are more capable of performing complex analytical processes. Currently, there are some notable research focused on this new area that demonstrate the value of integrating domain knowledge into visual analytics systems. Xiao et al. [1] presented a traffic analysis system to analyze network traffic using knowledge representation. Also Chang et al. [2] provided the city planning experts' insights through integrating the knowledge of urban legibility into geospatial visualization. Although these works have significance in helping reasoning and visualization, with few exceptions, the process of knowledge integration utilized in these systems is often specific to the analytical tasks or the domain, making it difficult to generalize to new problem areas.

Our approach focuses on representing data or information by determining the values of knowledge through understanding the data or information. Since finding and understanding important knowledge are extremely difficult, we begin with understanding the definition of knowledge and the user's analytical procedures. Specifically, we differentiate knowledge into tacit knowledge (personal, context specific, hard to formalize and communicate) and explicit knowledge (transmittable in a formal, systematic language) [6]. With this differentiation, we can formulate four knowledge conversion processes in knowledge-assisted vitalization: internalization, externalization, collaboration, and combination. A detailed explanation about each process is provided in section 3.

Based on the definition of knowledge and four knowledge conversion processes, we present a visual analytical system for bridge management for the US Department of Transportation. This system is visual in that it presents four important perspectives through visual representations: geospatial, temporal, relational and per-bridge detail. More importantly, it is directly connected to an ontological knowledge source and supports interactive data analysis to relate the user's domain knowledge to the ontology. We demonstrate that the four knowledge conversion processes of building (internalizing), generalizing (externalizing), sharing (collaborating), and merging (combining) knowledge help the analysis of bridge data by showing their utilities in two real-life scenarios.

The paper is arranged as the following. First we provide a detailed explanation of different definitions of knowledge (section 2). Based on our understanding of knowledge, we present the four knowledge conversion processes for knowledge-assisted visualization in section 3. In section 4, we demonstrate through an example how these processes could be incorporated in a visual analytical system. Finally we conclude this paper following a brief discussion.

2. Definition of Knowledge

To develop four knowledge conversion processes in knowledge-assisted visualization [7], we must first know what knowledge is. In the knowledge management literature, it has been established that distinguishing between data, information, and knowledge is important to designing knowledge management programs [8]. Work by Syed and Shah [5] reviews various definitions and explanations of the DIKW (data, information, knowledge, wisdom) hierarchy and focuses on presenting a model that explicates the relationship between data, information, and knowledge. In Syed and Shah's model, knowledge is defined as the range of one's information. However, Davenport and Prusak [9] state that "knowledge derives from information as information derives from data" and further define knowledge as "a fluid mix of framed experience, contextual information, values and expert insight that provides a framework for evaluating and incorporating new experiences and information." In Davenport and Prusak's perspective, knowledge is the refined information in which human cognition has added value. In other words, information becomes knowledge through cognitive effort.

Nonaka and Takeuchi [6] adopt Polanyi [3]'s definition of tacit and explicit knowledge to understand how knowledge is shaped and how knowledge can be applied. In their definition, explicit knowledge can be processed by a computer, transmitted electronically, or stored in a database. On the other hand, tacit knowledge is personal and specialized and can only be extracted by human. We extend Nonaka and Takeuchi's concept on knowledge conversion modes and apply them to visualization. We believe that through the use of interactive visualization tools, analysts can experience the interaction between tacit and explicit knowledge. To further delineate tacit and explicit knowledge in knowledge-assisted visualizations, we propose that:

- Explicit knowledge is different from data or information.
- Tacit knowledge can only result from human cognitive processing (reasoning).
- Explicit knowledge exists in data, and is independent from the user or his tacit knowledge.
- Explicit and tacit knowledge are related and can be connected through the use of interactive visualization tools.

Explicit knowledge, extracted from data or information, is represented as a visualization, which is received both perceptually and cognitively by the user via an image. The cognitive processing leads to an understanding and an increase of user tacit knowledge which recursively affects subsequent perception and cognition. Tacit knowledge guides the user's interaction and exploration so that the visualization changes over time. Based on the analytical expression of visualization proposed by van Wijk [11], tacit and explicit knowledge can be expressed as a set of equations:

$$K_e = f(D); I(t) = V(K_e, S, t); \frac{dK_t}{dt} = P(I, K_t) \quad (1)$$

where explicit knowledge K_e is a function and an extraction of data $f(D)$. Using explicit knowledge K_e, specifications S, an image at time t, $I(t)$, can be generated by the visualization system V. van Wijk [11] provides a detailed explanation about how the user interactively gains knowledge in using a visualization; the image I is perceived by P and understood by the user, resulting in an increase in the user's knowledge. Also he explains that the user may change the specification S of the visualization in order to explore E the data further. We believe that the user's gained knowledge is tacit K_t and depends on the image I and the current knowledge state K_t of the user. Since the amount of explicit knowledge K_e that exists in a complex dataset D could be nearly infinite, K_t can be expressed and stored in a knowledge base (KB) as a collection of smaller knowledge artifacts (K_{e_1} to K_{e_n}). Figure 1 shows a graphical representation of how explicit knowledge can be represented and where tacit knowledge might be located.

One important work that need to be recognized here is the knowledge-assisted visualization model proposed by Chen et al. [7]. While both Chen's model and our extended van Wijk model are fundamentally the same in structure and goal, our model differentiates knowledge into tacit and explicit forms. Since our focus is trying to distinguish and identify the four knowledge conversion processes in a visual analytics environment, we choose to express them based on extending the van Wijk's model, which shows a clear interrelationship among user, visualization and data.

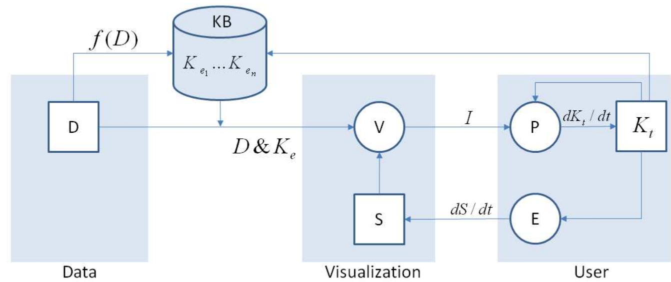


Figure 1: A graphical representation showing four entities: data, knowledge base (KB), visualization, and user. Once explicit knowledge is extracted from the data. It is stored in a knowledge base and used in visualization to represent it to a user. The user continuously perceives the image and gains tacit knowledge.

3. Knowledge Conversion Processes in Knowledge-assisted Visualization

Based on the proposed definitions of tacit and explicit knowledge, we provide four knowledge conversion processes in knowledge-assisted visualization: Internalization, Collaboration, Externalization, and Combination. These four knowledge conversions are first introduced by Nonaka and Takeuchi [6] in which they focus on how knowledge can be processed and converted from one to another in business models. For knowledge-assisted visualizations, we propose that the four processes are also applicable because the functions of an analysts in perceptually and cognitively understanding represented information to create concrete knowledge using a visualization is similar to that of analysts in business practices. Here we present the four conversion processes as applied to visualizations.

3.1. Internalization

In psychology, internalization is defined as the process of accepting the established set of norms, which are influential to the individual [12]. It is regarded as a cognitive process of acquiring skill and knowledge. In knowledge-assisted visualization, we propose that visually representing explicit knowledge would support analysts in understanding and transforming the explicit knowledge into tacit (internal) knowledge. As proposed by Nonaka and Takeuchi [6] the internalization process starts with a user discovering what the explicit knowledge is, followed by a series of steps in understanding why the explicit knowledge is of value or why it makes sense, until finally the user accepts the knowledge as their own viewpoint or internal knowledge (Figure 2). From a visualization perspective, this process parallels the concept of "insight discovery" that has been noted as the goal of visualization [13]. Since discovering insight is strongly related to building a user's tacit knowledge based on explicit knowledge in the data, the internalization process can be thought of as the primary goal and process of using a traditional (not knowledge-assisted) visualization.

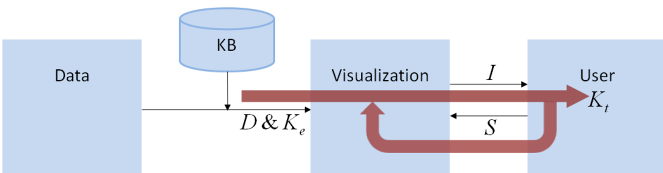


Figure 2: Internalization process (indicated by the red arrows). It explains that the user continuously builds tacit (internal) knowledge based on perceptually, cognitively, and interactively incorporating the represented explicit knowledge in a knowledge-assisted visualization.

3.2. Externalization

Externalization is the process of articulating tacit knowledge into explicit concepts [6]. It is a generic knowledge creation process through which tacit knowledge becomes explicit based on the user's finding (insights), concepts, hypotheses, and models. Since tacit knowledge cannot easily be shared with others as a result of simply being written down, it should be converted to explicit knowledge by an externalization process beforehand to communicate and share with others.

In the visualization community, there have been a few applications that specifically focus on the externalization of one's tacit knowledge. Garg et al. [15] presented a model-driven approach to extract Logic Programming (LP) rules through a user's interactions and reused the rules in further analysis processes. Xiao et al. [1] studied how the knowledge-base could be used to improve understanding of complex network traffic data. They found that about 80% of network traffic could be classified correctly based on previously extracted experts' knowledge-base. Chen et al. [7] showed an example in which the user's insights can be externalized into a knowledge base. These applications have shown that not only is externalization in a visualization possible, it is in fact a very powerful method for storing and reusing knowledge.

Figure 3 shows how tacit knowledge can be externalized into explicit knowledge. In this figure, explicit knowledge is stored in a knowledge base and used to assist the creations of visual representations. Because of the complex nature of explicit knowledge, Nonaka and Takeuchi suggest that externalization is the process of concept creation and is triggered by dialogue or collective reflection [6]. They further explain that the concept creation process can be expressed by applying analytical and non-analytical methods alike. The analytical methods are deduction and induction through which appropriate concept can be deduced and induced based on unorganized concepts. However, if applying analytical methods is not feasible, non-analytical methods such as metaphors and/or analogies could also be used. In visualization, we propose that both analytical and non-anlytical methods should be considered for expressing explicit knowledge that can in turn be stored into a knowledge base [1, 15].

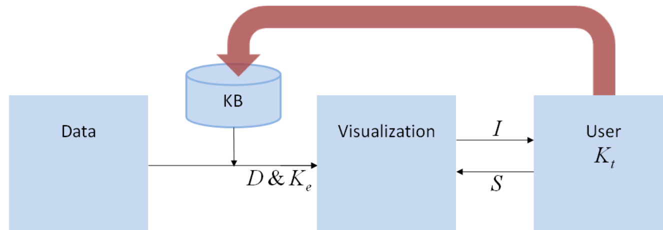


Figure 3: Externalization process (indicated by the red arrow). It shows that the internally created tacit knowledge can be extracted and saved into the knowledge base. The explicit knowledge can later be reused to create further visual representations.

3.3. Collaboration

Collaboration is the process of sharing tacit knowledge between people. In the knowledge management literature, it is defined as *socialization* [6]. Although both collaboration and socialization represent learning from others, we use the term *collaboration* because it is commonly used in computer science and has a history of implying sharing knowledge, learning, and building consensus through the use of computers. In visualization, building collaborative visualization environments also has a long history [16, 17]. Johnson [17] defined that collaborative visualization is a subset of computer-supported cooperative work (CSCW) in which control over parameters or products of the scientific visualization process is shared. Prior to that, Coleman et al. [16] provided generalizable reasons why collaborative visualization is compelling: (1) experts' knowledge could be available any time and at any place. (2) The expertise is transferred to others, improving the local level of knowledge. (3) Based on the supported accessibility, visualization products can be reviewed and modified as they are produced, reducing turn-around time. (4) Remote accessibility also helps to avoid relocating the expertise physically. More recently, Burkhard proposed a collaboration process of transferring knowledge between at least two persons or group of persons [4]. Similarly, Ma [18] noted that sharing visualization resources will provide the eventual support for a collaborative workspace. He discussed existing web-based collaborative workspaces in terms of sharing high-performance visualization facilities and visualizations and findings. He also showed several existing collaborative workspaces such as TeraGrid [19], Many Eyes [20], etc (see [18] for detail).

Figure 4 shows how two visualization users could collaborate and share their tacit knowledge with each other. In this diagram, we show that the collaboration process can occur through the use of collaborated visual environments. However, the most natural method for sharing knowledge is still direct communication between the users (via phone, email, instant messages, etc.). In either case, the users are actively sharing their discoveries and tacit knowledge and incorporating each other's domain expertise into their own.

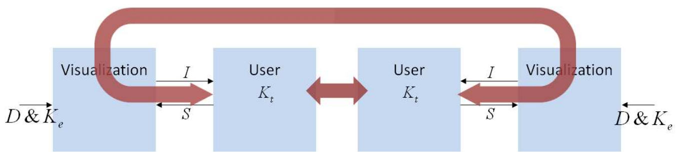


Figure 4: Collaboration process (indicated by the red arrows). Collaboration is a process of sharing tacit knowledge through the use of a visualization or through direct communication.

3.4. Combination

Combination is the process of systemizing explicit concepts into an explicit knowledge system [6]. Since explicit knowledge exists everywhere in books, research papers, and communication networks (user groups), etc., the process of combining different bodies of explicit knowledge is important. For instance, genomic data have been used in many different research areas: biology, bioinformatics, computer science (including visualization), health & medical science, etc., and depending on the domain, researchers have derived different, yet equally important findings. In order to fully comprehend the knowledge associated with such genomic data, it is necessary to combine findings from different domains and integrate them into a cohesive set of explicit knowledge.

Figure 5 shows a simple model of combining explicit knowledge into an existing knowledge-base. Behind this simple diagram, however, additional considerations need to be addressed in order to maintain the quality and integrity of the knowledge-base when combining new explicit knowledge with an existing source. If unrelated or incorrect knowledge is combined with the existing explicit knowledge, it could degrade the overall trustworthiness of the knowledge-base as well as the benefits of representing knowledge in a visualization. While we are not aware of any known visualizations that support the verification and validation process when combining explicit knowledge, in the Knowledge Engineering literature, researchers have long been studying how to verify and validate underlying knowledge when developing a knowledge-based system [21]. The specifics of knowledge management and engineering is outside the scope of this section, but will be discussed further in the future work section (section 5).

3.5. Associating Knowledge Conversion Processes

We define the four processes described above based on understanding of the existing knowledge conversion modes [6]. We believe that these four knowledge conversion processes are also important for knowledge-assisted visualization in that they can be utilized to creating more effective visual analytical systems. Internalization is the process of gaining tacit knowledge from continuous interactive analysis and is exemplified

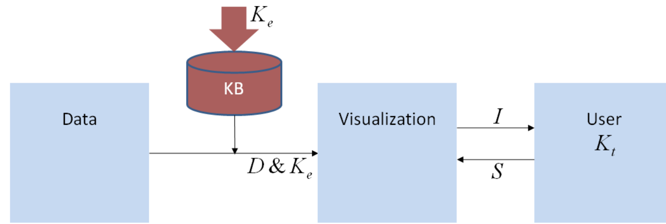


Figure 5: Combination process (indicated by the red arrow). Explicit knowledge can be combined with existing explicit knowledge in a knowledge base to further enhance a user's visual analysis process.

by the human's reasoning process of finding insights in visualization. Externalization is the process of generalizing the found insights, concepts, hypotheses, and models during the use of a visualization into externalized forms. Both analytical and non-analytical methods can be used to express the findings. Collaboration occurs when users share discovered knowledge from analyzing data in visualization, and it represents an important facet of knowledge visualization in which a user's finding is directly benefiting others. Finally, the users' findings should be generalized and integrated into existing knowlege-bases for future use through the combining process.

Although the four knowledge conversion processes seem independent of each other, complex analysis tasks often utilize two or more of these processes. For instance, both internalization (user acquiring tacit knowledge from information displayed on the screen) and externalization (user making annotations that become transmittable explicit knowledge) are processes that take place in most investigative analyses. Since a visualization system could be used to establish a two-way communication between a user and a knowledge structure by providing highly interactive and exploratory capabilities, we believe it could be an environment in which all four knowledge conversion processes are integrated and utilized. However, to the best of our knowledge, there does not exist any visualization that incorporates all four knowledge conversion processes, but we believe that each process represents an important area of knowledge-assisted visualization and should be considered together in a holistic way.

4. Applying Knowledge Conversion Processes into Visualization

As shown in Figure 1, a well-designed knowledge base plays an important role in supporting the knowledge internalization, externalization, collaboration, and combination processes. We believe that in order to design a useful visual analytics system that incorporates knowledge, a tightly integrated and well-designed knowledge base is essential.

There is, however, no definitive way to construct a knowledge base. Much research has focused on designing and developing different forms of such databases that could represent domain knowledge. The differences between these database are not only reflected in their capacities, but also in their structural complexities. As shown in work by Garg et al. [15], a knowledge base could be as simple as a textual structure that contains inductive logic programming equations. On the other hand, it could also be described by extensive decision models, such as Markov decision process (MDP) in the Artificial Intelligence field. In our example, we choose to apply an ontology for storing and retrieving domain specific knowledge.

The ontological knowledge structure is a conceptualization of domain knowledge which includes concepts, properties and their relationships. This conceptualization process aims to transfer both human tacit knowledge and explicit knowledge into computer-understandable formats. These concepts can be further utilized to facilitate other users' problem-solving processes. More specifically, a Problem Domain Ontology (PDO) enables solving a complex problem where the underlying domain concepts have high interdependencies by building up a problem scenario based on concepts, properties and features in the ontological knowledge structure.

Although research on ontological knowledge structure have advanced in the recent years, integrating such structure with a visual analytics system is still an open research area. In the following subsections, we first describe our understanding about how to integrate these two components, and further present our prototype of a knowledge-assisted visual analytics system.

4.1. *System Overview*

4.1.1. *The Relationship between Visualization and Ontology*

While there is no definitive guideline on how to integrate visual analytics with ontological knowledge structure, we hypothesize that the integration of these two approaches could form a useful knowledge-assisted visual analytics environment. In order to have a better understanding of why this integration is meaningful and feasible, we examine visual analytics and ontological knowledge structure separately for their capabilities and strengths. While visual analytics usually allows the user to interactively explore patterns of the underlying data from various perspectives, the ontological knowledge structure focuses more on representing the conceptualization of domain knowledge and the interdependencies among the concepts. Although through distinctive approaches, both visual analytics and ontological knowledge structure help the user to understand and discover different aspects of knowledge.

If these different approaches could be reasonably integrated, we expect that users could discover new concepts and knowledge through exploring the visualization and externalize such knowledge into the ontological knowledge structure for future references. We also want users to directly access the knowledge structure to acquire predefined domain concepts and rules to guide them through visual explorations and assist their decision-making processes.

To validate our hypothesis, we present our bridge management system which utilize encapsulated knowledge from a domain specific ontological knowledge structure. Further, we demonstrate two scenarios to show where such integrated system could be deployed.

4.1.2. *System Implementation*

With support from the US Department of Transportation (USDOT), we implemented a prototype of knowledge-

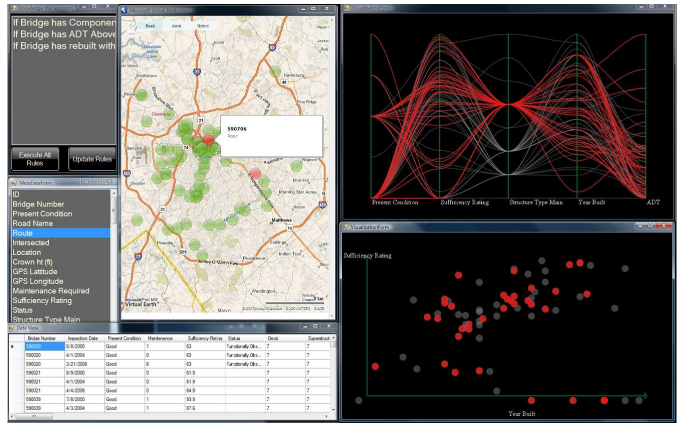


Figure 6: The overview of the knowledge-assisted visualization system. Top Left: the Knowledge Window. Middle left: the meta data window. Middle: The Geospatial View. Top Right: the Parallel Coordinate View. Bottom Right: the Scatter Plot View.

integrated visual analytics bridge management system. Our system contains two major components: a visualization interface that provides interactive data exploration, and an ontological knowledge structure that is customized to store and provide bridge management domain concepts and knowledge. Through cyclic communications between these two components, our system provides bridge managers with a comprehensive understanding about their bridge assets and facilitates their decision-making processes.

Visualization Interface. As shown in Figure 6, our system utilizes a highly interactive visualization interface to help depict bridge data from three aspects: geospatial, temporal, and relational. Utilizing the rich geo-information provided by Microsoft Virtual Earth [24], our system provides an interactive geospatial view for the bridge managers to examine the distribution of bridges as well as their surrounding environments. To enable temporal analysis, we designed a Treemap-based [22] small multiples [23] view to represent the temporal trends of individual bridges with a spatial layout generated based on user-chosen dimensions. The Parallel Coordinate [25] and Scatter Plot views are dedicated to assist bridge managers in depicting the relational information among bridges and their attributes. By tightly coordinating these views together, our system provides the bridge managers an interactive data exploration environment that could help them comprehend complex bridge information from multiple perspectives simultaneously.

The Ontological Knowledge Structure. In addition, an ontological knowledge structure is integrated into our system to provide domain concepts and information. Using an ontology-driven modeling approach [26, 27], this ontological knowledge structure contains bridge domain concepts, such as bridge structural types and locations. These individual bridge concepts are further connected through their interdependent relationships, which is modeled based on the experience of bridge managers and other domain users. By connecting concepts in such a manner, additional domain rules can be identified and created.

For example, a rule, which suggests the bridge would have potentially undergone severe structural damage, can be described as: if a bridge's sufficiency rating is below 50 and its superstructure rating is less than 5. Such rule would be created into the knowledge structure and will be executed to alert users the situation upon requests. Utilizing such a rule-based ontological knowledge structure allows for great flexibility for our system to support precise examination of bridges and enables the system to better facilitate bridge management processes.

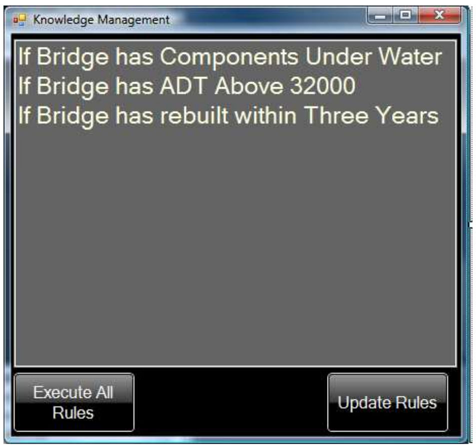


Figure 7: The knowledge window provides updated knowledge rules inside the ontological knowledge structure.

Communication between components. Through a server-client web interface, our system tightly coordinates the visualization interface with the ontological knowledge structure. Since these two components share the same underlying bridge ID number, the message passing becomes clear and feasible. For example, any results from the executed rules in the ontological knowledge structure will be immediately updated in each visualization window. Thus, exploring within visualization could lead to new concepts that can be further added into the ontological structure; while the knowledge stored in the ontology could assist decision-making during the visual exploration.

In order to assist bridge managers in executing the domain rules, our system presents an interactive knowledge window (Figure 7) which is automatically synchronized with rules within the ontological knowledge structure. With these two components tightly integrated together, the users always have access to the most up-to-date rules and concepts. The users simply have to execute the relevant rules, and they can see and interact with the bridges in detail immediately in the visualization environment.

Furthermore, our system enables the bridge managers to directly modify the knowledge structure. This function provides bridge managers an important interface to update the externalized knowledge and maintain its accuracy. Based on their discoveries during their interactions with the visualization, bridge managers could create new concepts or rules and directly insert them into the ontological knowledge structure. For example, through their interaction with visualization, bridge managers may find that the combination of low ratings (less than 4) on both "supporting structure" and "water adequacy" suggests water erosion and flood damage. The bridge managers could then insert this new discovery into the ontological knowledge structure and further re-apply it to check how many bridges have been affected by water-erosion or damage.

Embedded Knowledge Processes. Since this bridge management system is designed based on our definition of knowledge and its corresponding conversion processes, we can clearly identify the four different knowledge conversion processes - internalization, externalization, collaboration, and combination in its functions:

- The Internalization process embodies the transfer of knowledge from a computer to a user through the interactions with a visualization. In our system, this process mainly happens through the user's interaction with the coordinated visual analytics views. These views help the users inspect the data from different perspectives and assist the potential discovery of unexpected data patterns and trends that could become new domain knowledge.

- The Externalization process happens upon the user's acquisition of new domain knowledge or information that does not already exist in the ontological knowledge structure. This knowledge could come from both discoveries from interacting with the visualization system or from collaborating with other co-workers. Once acquired, the user could directly insert this new knowledge into the ontological knowledge structure to augment its knowledge base. The ontology will then store this knowledge and re-apply it during a user's future investigations.

- The Collaboration process takes place when a user interacts with our integrated system that incorporates domain knowledge of multiple experts. Through our integrated knowledge interface, each bridge managers connects to the same ontological knowledge structure. New knowledge or domain rules created by one manager would immediately be reflected in another bridge manager's visualization system. In this manner, through the use of the ontology as a central repository of knowledge, our system facilitates collaboration between multiple bridge managers.

- The Combination process occurs when inserting new knowledge into the existing knowledge structure. The new knowledge could come from a new set of domain data, new perspectives or regulations on bridge inspections, etc. Since bridge inspection rules vary for different inspection cycles due to new federal bridge inspection guidelines or regulations, the Combination process is particularly important in ensuring that each bridge manager is inspecting their data with the most suitable domain knowledge. For example, to handle changes in the standards of water adequacy, our system combines different sets of that criteria and applies them accordingly to different inspection cycles.

4.2. Scenarios

We performed a qualitative (expert-based) evaluation of our system with bridge managers from North Carolina and City of Charlotte Department of Transportation and identified the following scenarios that demonstrate how our knowledge-assisted visualization system could assist bridge managers' daily jobs of examining bridges and making maintenance decisions.

4.2.1. Augmenting Visualization through the use of an Ontology

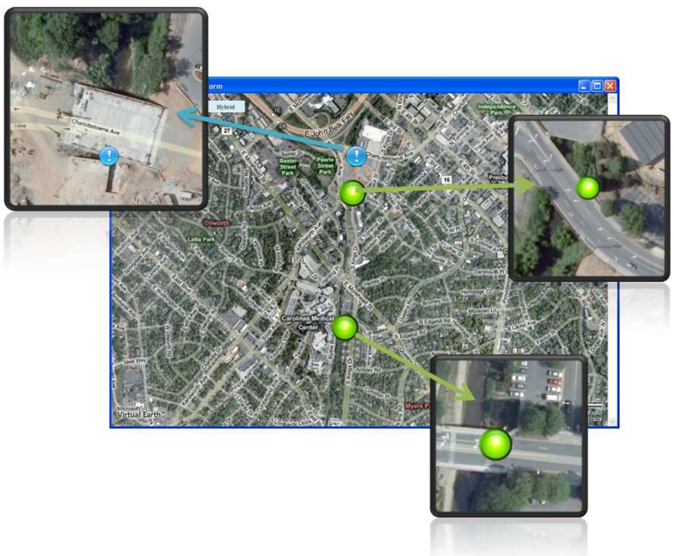


Figure 8: Close examination of the geospatial view shows that although these three bridges are on the same river stream, their conditions are different. The bridge over the upper stream is currently under repair and reconstruction.

According to bridge experts, water erosion and flooding can cause severe damages to bridges. The pattern for this type of deterioration is in general typical along river streams. In this scenario, we demonstrate how our system could help bridge managers to quickly identify the cause of unexpected bridge deteriorations through the knowledge internalization process. This scenario was identified together with city of Charlotte bridge management team during their examination of causes of water damage.

Since the criterion for "bridge above water" has already been externalized in our ontological knowledge structure, the bridge managers can easily highlight all these bridges in the geospatial view and examine them individually. Through quick examination on the geospatial view, the bridge managers immediately noticed an interesting pattern in South Charlotte. Although located over the same river, as shown in Figure 8, the three bridges over that river showed different "present conditions". The one over the upper stream has already been filed for replacement and has been under construction. However, the other two are still in good condition. This pattern is interesting because if there was a flood, all three bridges should share similar deterioration patterns; or at least, they should deteriorate at a similar pace. Even though temporal information suggests that these bridges were built at similar times, the changes in their conditions are drastically different. This inconsistency raised the bridge managers' interests.

After a detailed examination of these bridges in the geospatial view, the bridge managers realized that the cause of this inconsistency was due to the different turns of the river. According to one of the bridge managers, although there was flood in both the upper and lower parts of this river, the bridge over the upper stream received the most impact since there were no bends in the river before the water hit it. On the other hand, due to the slow down of the river's speed when the water passed the second and third bridges, these two bridges received much less impact. Based on this observation, the bridge management team was able to quickly identify and internalize this pattern and re-use it for future reference.

In this scenario, the bridge managers gained insightful knowledge from interacting with our visualization system and incorporated it into their tacit knowledge (internalization). Although it would be more beneficial and efficient if this kind of knowledge could be externalized, due to the complexity of modeling such knowledge, our current ontological knowledge structure does not support an explicit externalization for it.

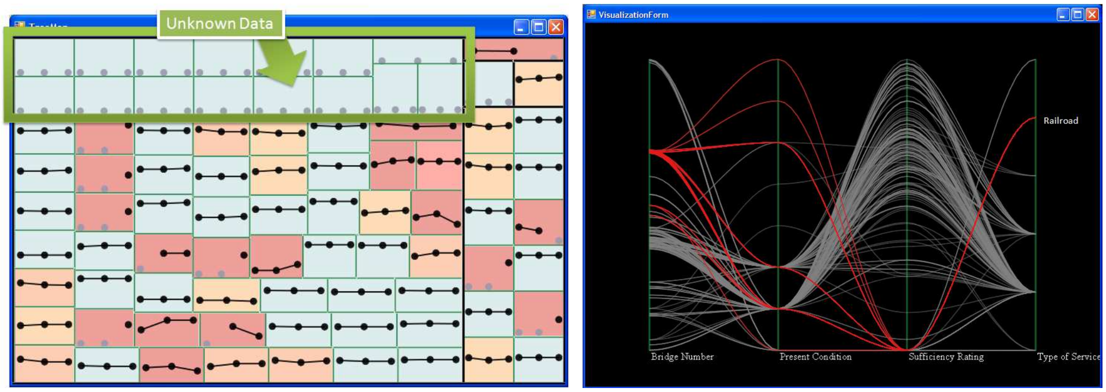
4.2.2. Updating and Sharing Knowledge through Visualization

Since managing bridges is a complex process that often requires precise analysis, it is important for a bridge analyst to quickly determine the most relevant information to focus on during an investigation. In this scenario, we demonstrate how our system facilitates bridge experts through the externalization of their discoveries and sharing the findings (collaboration) to filter out unnecessary data and focus on analyzing the most relevant information.

A local bridge expert was using our visualization system to explore the bridge distributions around the Charlotte region. After a quick examination of the temporal view, the expert noticed that a large group of bridges did not have any ratings information (Figure 9(a)). Based on this bridge expert's experience, this situation was most likely caused by two reasons: one, it could be caused by a loss of data or errors during the data entry process. Two, these bridges could be outside of the bridge management team's jurisdiction. As shown in the coordinated visualization views (Figure 9(b)), the bridge experts identified that these bridges were all railroad bridges, which fell into the second category.

Since the city bridge management team is not responsible for maintaining these bridges, showing them together with other bridges can be confusing. In order to reduce this confusion, the bridge manager created new rules in the ontological knowledge structure to identify and filter out these railroad bridges. Other bridge managers of the same team will then be able to reuse these rules to reduce the irrelevant information and concentrate on the relevant bridges.

This scenario shows how a user could gather information during visual exploration and further update (externalization) and share his knowledge discoveries with other co-workers (collaboration).



(a) (b)

Figure 9: (a) A large group of unknown data is shown in the temporal view, which lead to the search of its cause. (b) Visualization views indicating that these are Railroad bridges.

5. Future Work and Discussion

Although the four knowledge processes are shown to be effective in transferring and storing knowledge between multiple users in our bridge visualization system, there are still several remaining issues that need to be solved, including one fundamental challenge on how to verify and validate newly inserted knowledge. This challenge, through our discussion with domain experts, has its significance in determining the usefulness of a knowledge-assisted visualization system.

As mentioned in the Combination process, if new knowledge is not carefully validated, inserting unrelated or incorrect knowledge could potentially degrade the value of the ontological knowledge structure. However, verifying and validating diverse human knowledge is difficult in nature. Due to individual experiences and understanding, different experts have their own ways of representing and describing knowledge. In general, there are four potential issues in inserting newly created knowledge into an existing knowledge base: duplicated, partial-overlapped, imprecise and conflicting knowledge.

- Duplicated knowledge is often created when different experts share the same discoveries without realizing that one already exists. For example, bridge managers often apply the rule that regulates bridges with sufficiency-rating value below 50 and super-structure rating below 6 in order to quickly decide which bridges require further examination. However, depending on the size of the knowledge base, an expert user might not know that this rule already exists and creates an identical one (but perhaps represented slightly differently using a different grammar or language).

- Partially overlapped knowledge in the knowledge base represents rules that are mostly similar but contains subtle yet important differences. As in the previous example, some bridge managers may like to add the substructure value into consideration. Since it is not simple to decide if the newly added rules should be inserted or merged with the existing one, verifying this type of relationship is difficult. One conservative approach could be to provide multiple copies of the overlapped rules and group them together for individual users to choose from.

- Imprecise knowledge would be introduced into the knowledge base while the knowledge itself is still developing. This situation frequently occurs during the bridge management process, where new patterns and issues keep emerging. For example, while one crack on the east side of a bridge should not draw much concern on the overall structure of the bridge, a initial knowledge rule describing this situation could be imprecise, considering that it does not specify the size or width of the crack. One way to validate and improve the use of this kind of knowledge is to apply interactive methods between the knowledge base and users. Through these interactions, users could insert more precise knowledge to perfect the associated rules and provide more accurate future reasoning references.

- Conflicting knowledge occurs when existing knowledge differs from newly created knowledge due to perhaps new protocols or regulations. Unfortunately, without the supervision of domain experts, most systems cannot effectively identify and resolve such conflicts. This is particularly challenging when the existing knowledge is still valid (and should be kept in the knowledge-base) for decisions made in the past or under specialized circumstances (e.g., a drastic change in the budget). Due to the multiple possible causes of conflicted knowledge, organizing them is arguably the most difficult and least understood.

Understanding each of these relationships is imperative for maintaining the validity and integrity of a knowledge base that is used in real decision-making environments. At present time, all three scenarios are handled by domain experts manually. However, in the Knowledge Engineering literature, researchers proposed and designed several verification and validation (V&V) techniques and tools [21]. Some of them support the ability to automatically verify and validate underlying knowledge. But without a clear understanding of domain knowledge, most automatic techniques and tools are not always reliable. In knowledge-assisted visualization, it remains an open research area for us to create a semi-automated knowledge management system for organizing and storing diverse knowledge and rules in the same knowledge base.

6. Conclusion

In this paper, we first propose our definition of knowledge in the context of visual analytics and further differentiate knowledge into two types, tacit and explicit. With the differentiation, we then formulated four knowledge conversion processes in knowledge-assisted visualization: internalization, externalization, collaboration, and combination. Based on the definition of knowledge and the four knowledge processes, we designed a visual analytics system for bridge management with the support from US Department of Transportation. By connecting to an ontological knowledge source, our visual analytics system allows users to interactively analyze the data with access to the expert's domain knowledge. We demonstrate through this bridge management visualization system that the four knowledge conversion processes are feasible and applicable to visualization systems and should be considered together when designing knowledge-assisted visualizations that incorporate domain expert knowledge.

References

[1] Xiao L, Gerth J, Hanrahan P. Enhancing Visual Analysis of Network Traffic Using a Knowledge Representation. In: Proceedings of Visual Analytics Science And Technology 2006. IEEE. p.107-114.

[2] Chang R, Butkiewicz T, Ziemkiewicz C, Wartell Z, Ribarsky W, Pollard N. Legible Simplification of Textured Urban Models. IEEE Computer Graphics and Applications 2008;28(3):27-36.

[3] Michael Polanyi, Tacit Dimension, Peter Smith Publisher Inc, 1983.

[4] Burkhard RA. Learning from architects: the difference between knowledge visualization and information visualization. In: Proceedings of the 8th International Conference on Information Visualization 2004. IEEE. p.519-524.

[5] Syed A, Shah A. Data, information, knowledge, wisdom: A doubly linked chain? In: Proceedings of the 101In International Conference on Information and Knowledge Engineering, pages 270-278, 2006.

[6] Nonaka I, Takeuchi H. The Knowledge Creating Company. Oxford University Press; 1995.

[7] Chen M, Ebert D, Hagen H, Laramee RS, van Liere R, Ma K, Ribarsky W, Scheuermann G, Silver D. Data. Information, and Knowledge in Visualization. IEEE Computer Graphics and Applications 2009;29(1):12-19.

[8] Jurisica I, Wigle D. Knowledge Discovery in Proteomics. 2nd ed. CRC Press; 2005.

[9] Davenport TH, Prusak L. Working Knowledge: How Organizations Manage What They Know. Harvard Business School Press; 1998.

[10] The American Heritage Medical Dictionary 2007.

[11] vanWijk JJ. The value of visualization. In: Proceedings of IEEE Visualization 2005. p.79-86.

[12] Meissner WW. Internalization in Psychoanalysis. New York: International Universities Press; 1981.

[13] Chang R, Ziemkiewicz C, Green TM, Ribarsky W. Defining Insight for Visual Analytics. IEEE Computer Graphics and Applications 2009;29(2):14-17.

[14] Yi JS, Kang Y, Stasko JT, Jacko JA. Understanding and characterizing insights: how do people gain insights using information visualization?. In: Proceedings of the 2008 Conference on Beyond Time and Errors: Novel Evaluation Methods For information Visualization, Florence, Italy, April 2008.

[15] Garg S, Nam JE, Ramakrishnan IV, Mueller K. Model-driven Visual Analytics. In: Proceedings of Visual Analytics Science and Technology 2008. Oct. 19-26. p.19-26.

[16] Coleman J, Goettsch A, Savchenko A, Kollmann H, Wang K, Klement E, Bono P. TeleInViVo^{TM}: Towards collaborative volume visualization environments. Computer & Graphics 1996;20(6):801-811.

[17] Johnson G. Collaborative visualization 101. Computer Graphics 1998;32(2):8-11.

[18] Ma K. Creating a collaborative space to share data, visualization, and knowledge. SIGGRAPH Comput. Graph. 2007;41(4).

[19] Binns J, DiCarlo J, Insley JA, Leggett T, Lueninghoener C, Navarro J-P, Papka ME. Enabling community access to TeraGrid visualization resources. Concurrency and Computation: Practice and Experience 2007;19(6):783-794.

[20] Many Eyes, http://manyeyes.alphaworks.ibm.com/manyeyes.

[21] Tsai W, Vishnuvajjala R, Zhang D. Verification and Validation of Knowledge-Based Systems. IEEE Trans. on Knowl. and Data Eng. 1999;11(1):202-212.

[22] Bruls M, Huizing K, Van Wijk JJ. Squarified Treemaps. In: Proceedings of the Joint Eurographics and IEEE TCVG Symposium on Visualization 2000, Eindhoven University of Technology p.33-42.

[23] Tufte ER. Envisioning Information. Graphics Press; 1990.

[24] Microsoft Corporation, http://www.microsoft.com/virtualearth.

[25] Inselberg A. The plane with parallel coordinates. The Visual Computer 1985;1(4):69-91.

[26] Lee SW, Gandhi RA. Ontology-based Active Requirements Engineering Framework. In: Proceedings of the 12th Asia-Pacific Soft. Engg Conference 2005, IEEE CS, p.481-490.

[27] Lee SW, Muthurajan D, Gandhi RA, et al. Building decision support problem domain ontology from natural language requirements for software assurance. International Journal on Software Engineering and Knowledge Engineering 2006;16(6):1-34.