Knowledge-assisted Visualization in the Cultural Heritage Domain
Case Studies, Needs and Reflections

Patricia Martín-Rodilla
Institute of Heritage Sciences (Incipit), Spanish National Research Council (CSIC),
Rúa San Roque, 2, Santiago de Compostela (A Coruña), Spain

Keywords: Knowledge-assisted Visualization, Discourse Analysis, Cultural Heritage, Inferences, Visualization Requirements.

Abstract: We investigate on how to build software systems that assist cultural heritage researchers to reconstruct past events on the basis of present data. In this setting, knowledge-assisted visualization can be a useful mechanism to improve the knowledge generation process and emphasise collaboration. However, a useful visualization depends on the goals of the user and the specific research problem involved. In this position paper we present a set of case studies to defend the study of cognitive inferences through discourse analysis and its typologies as a starting point in the knowledge-assisted elicitation process. Once a complete study of usual inferences in the cultural heritage domain is done, the visualization needs in this domain will be easier to determine and apply, attaining our objective of knowledge-assisted visualization.

1 INTRODUCTION

Cultural heritage activities such as archaeological excavations or anthropological studies manage a large amount of complex data. In previous empirical studies, an elicitation process about the interaction preferences of researchers in cultural heritage (Martín-Rodilla, 2012) or conceptual modelling studies in cultural heritage domain (González-Pérez and Parcero-Oubiña, 2011) were carried out as a previous work to understand the needs of cultural heritage researchers. The visualization of this data and it application to generate knowledge is one of the problems detected in these studies.

Some literature and applications of visualization in cultural heritage exist (Vote et al., 2002); (De Luca et al., 2011). However, these approaches are, in most cases, *ad hoc* proposals not aimed at knowledge-assisted visualization.

Knowledge-assisted visualization involves a number of areas in visualization theory, including theory frameworks and taxonomies (Zhou and Feiner, 1998); (Amar and Stasko, 2005), cognitive studies about the knowledge generation method and visualization cognition (Huang et al., 2006), data characterization (González-Pérez, 2012) or interaction principles (Van der Vlist el al., 2011).

This paper shows an initial study based on discourse analysis as technique to characterize the usual inferences that occur in the cultural heritage domain and their implications in knowledge-assisted visualization.

2 METHODOLOGY AND PROPOSAL

According to the literature (Tufte, 1990); (Chen et al., 2009) it is possible to assist in the knowledge generation process of humans by improving the visualization of the data and the processes involved.In knowledge-assisted visualization (Chen et al., 2009), the user's knowledge is an important part of the visualization. Characterizing this knowledge, we can incorporate it as expert knowledge, building best-fit visualizations in a specific domain.

The question that we find unanswered, therefore, is what are the visualization needs in the cultural heritage domain to design useful knowledge-assisted visualizations and what is the best method to characterize this knowledge.

We propose a deep study of the most common types of inferences, in order to correctly support the processes of knowledge generation in a specific

Martín-Rodilla P..
Knowledge-assisted Visualization in the Cultural Heritage Domain - Case Studies, Needs and Reflections.
DOI: 10.5220/0004281605460549
In *Proceedings of the International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications* (IVAPP-2013), pages 546-549
ISBN: 978-989-8565-46-4

domain. The future users of a visualization tool or technique in this field are archaeologists, anthropologist, historians and researchers or practitioners in similar disciplines. These professionals aim to reconstruct past events through present data. For this reason, the best-fit visualizations to assist the processes that they carry out are those that can support the most common inferences that occur in the cultural heritage domain, which often entail strong spatio-temporal components.

The case studies presented below were developed by applying the technique of discourse analysis (Hobbs, 1985) in order to dissect the relevant cultural heritage statements and determine visualization user needs. This method analyses the syntactic of the discourse looking for elements of connection such as prepositions or similar particles and evaluates the knowledge implied in the text. Hobbs's analysis is put into practice in other research areas such as biomedical domain (Lacson et al., 2006), legal arguments (Moens et al., 2007) or mathematics research (McGrath and Kuteeva, 2012) to extract information and requirements. We try to find common inferences types in archaeological research as a subfield of cultural heritage. The differences between archaeology and other cultural heritage disciplines in this argumentation and the generalization of the conclusions are future needs in our research.

Establishing types of inference presented in the discourse allows us to determine the needs of the users, introduce it as expert knowledge and elaborate a set of suggested visualization solutions that could be used as input for a research-oriented software system.

3 CASE STUDIES AND NEEDS

The following sections present three case studies. Each case study includes a literal statement taken verbatim from a research work in archaeology, an analysis of the inference according with the characteristics of the discourse in said statement, and a reflection about the user needs that this analysis reveals in terms of visualization requirements and suggested visualization techniques.

3.1 Case Study One

- Cultural Heritage Statement: "There is a group of pots with no decoration, low quality clay and hand made. They could have been used as housewares. This hypothesis is fitting with the chemical analysis of the pots" (Cobas and Parcero-Oubiña, 2006).
- Inference Analysis: This is a common kind of hypothesis in archaeology. Based on the classification of pots or pot fragments found at a particular site, researchers assess the quality and analyse the chemical properties of the pots, and they try to reconstruct the function they had. The first objective is to find groups or similarities between pots. The presence or absence of decoration, the quality of the clay, and the manufacturing process (such as by hand or wheel) for each pot are determined and compared with the outcomes of the chemical analysis.
- User needs: For this kind of inference, the first process is one of clustering. The second process is a relational one: the researcher wants to find a relation between the pot groups and the characteristics of the pots in order to infer the function or common use of the pots. To improve this knowledge-generation process, the researcher needs to visualize data in groups. To support the second process, the researcher may need to focus on a particular group and visualize the specific characteristics of its pots and their chemical properties in order to determine the function of the pots.

3.2 Case Study Two

- Cultural Heritage Statement: "In funerary contexts in the Bronze Age of Atlantic Europe, we can find individual tombs with decorated and high-quality manufactured pots with handles and brilliant colours" (Prieto-Martínez, 1998).
- Inference Analysis: In this case, the first part of the statement is a selection of a space-time slice. The researcher relates the funerary pots with their characteristics (material, colour, quality, handles). The main objective is the characterization of the funerary pots in Atlantic Europe during the Bronze Age.
- User needs: In order to compare the target pots with pots of similar but different contexts, the researcher needs to visualize a map of their spatio-temporal situation plus the characteristics being analysed (material, colour, quality, handles) for each region of the map. In addition, the researcher may want to visualize the sequence of temporal events for a

specific area of the map, or for a specific pot, or other spatio-temporal relationships.

3.3 Case Study Three

- Cultural Heritage Statement: "The source of inorganic matter implies that the scope of interaction does not exceed 50 km. This means that animals were required to transport the wheat to the mills" (Fernandez, 2010).
- Inference Analysis: The first part of the inference involves the grouping of inorganic matter found, as well as the delimitation of the area where those inorganic matters were found. Knowing that traces of wheat farming were found on the site, the research infers that animals must have been required to transport the wheat, because no mills were present within the studied archaeological area.
- User needs: This kind of inference is more complex than the previous two, because it involves a concatenation of simpler inferences. In the first part, the researcher needs to visualize the area where inorganic matter and wheat farming traces were found. In the second part, the researcher needs a relational visualization to match the distance and the nearest mills on the map.

3.4 Overall Observations

There are some differences in the inference types and knowledge-generation methods that are revealed in the case studies presented. These differences mean that the knowledge-assisted visualization requirements are arguably different too.

In addition, we detected that the most common types of inferences are clustering or relational mechanisms. This idea is not exclusive in cultural heritage and it is applicable to other research areas (Morato et al., 2003); (Fujita and Teplovs, 2009). However, this characterization allows us to identify a set of needs in visualization for researchers in cultural heritage and suggest some association visualization techniques:

- We detect clustering inference as a starting step in the knowledge-generation process. We think that cluster visualization and related techniques (Van Wijk and Van de Wetering , 1999); (Stasko and Zang, 2000) are good choices for the first step in the knowledge-assisted process.
- We detect strong spatio-temporal components in the following steps, e.g. in cases two and three. It is arguable that techniques related to spatio-temporal visualizations (Andrienko et al., 2003); (Aigner et al., 2011) are good candidates for the next steps in the knowledge-assisted process.
- We also detect needs in terms of varying granularity of the information that it showed in the visualization. In all three case studies, statements involve large amounts of data. In this situations we argue that the application of well-known patterns such as Abstract/Elaborate (Yi et al., 2007), Filter (Yi et al., 2007), Snap-Together (North and Schneiderman, 2000), or patterns at the implementation level (Heer and Maneesh, 2006) would work well to solve these needs.

This list of needs is not complete. The main purpose of this paper is to start a discussion about the analysis of discourse and the typology of inferences as a method to understand a domain and, in particular, the cultural heritage domain, and improve knowledge-assisted visualization.

4 REFLECTIONS AND DISCUSSION

This study is an initial conceptualization about the main features in the knowledge-generation process in cultural heritage and it needs about visualizations.

Our final objective it is to introduce this domain between the topics in visualization, to discuss scientifically the needs extracted in this study and compare with other domains with a successfully application of knowledge-assisted visualization techniques.

In this paper we defend the need of a deep study about the knowledge-generation process in each domain in terms of inferences and discourse analysis, as a part of the process of elicitation requirements in the knowledge-assisted visualization.

This proposal is an appropriate initial-point in the requirements visualization process, because allow us to understand the mind process involved in each domain and the final objectives of each research analysed. With this information, we could choose better visualization principles and techniques for each objective. Following these ideas, our future work is focused on the conceptualization of the method and in the elaboration of test-bed and empirical studies with users.

REFERENCES

Aigner, W.; Miksch, S.; Schumann, H. and Tominski, C. (2011).*Visualization of Time-Oriented Data*, Springer.

Amar, Robert A.; Stasko, John T. (2005). Knowledge Precepts for Design and Evaluation of Information Visualizations. In *IEEE Transactions on Visualization and Computer Graphics* 11(4): 432 – 442.

Andrienko, N., Andrienko, G. et al. (2003). "Exploratory spatio-temporal visualization: an analytical review." In *Journal of Visual Languages & Computing* 14(6): 503-541.

Chen, M.; Ebert, D.; Hagen, H. et al. (2009). Data, Information, and Knowledge in Visualization. In *IEEE Computer Graphics and Applications* 29(1): 12-19.

Cobas, I.; Parcero-Oubiña, C. (2006). Alto Do Castro (Cuntis, Pontevedra). Síntesis De Resultados Y Estudios De Materiales, Campaña 1993. In *Tapa (Traballos De Arqueoloxía E Patrimonio)*, Spanish National Research Council. Santiago de Compostela.

De Luca, L., Busayarat, C., Stefani, C., Véron, P., Florenzano, M. (2011). A Semantic-Based Platform for the Digital Analysis of Architectural Heritage. *Computers and Graphics* 35(2): 227-41.

Fernández Martín, S. (2010). Los complejos cerámicos del yacimiento arqueológico de La Motilla del Azuer, PhD Thesis. University of Granada.

Fujita, N.; Teplovs, C. (2009). Automating analysis of collaborative discourse: Identifying idea clusters. In C. O'Malley, D. Suthers, P. Reimann & A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices.* International Society of the Learning Sciences, Inc. Rhodes, Greece: 162-164.

González-Pérez, C. (2012). A conceptual modelling language for the humanities and social sciences. In *RCIS'12: Sixth International Conference on Research Challenges in Information Science*, Valencia, Spain.

González-Pérez, C., Parcero-Oubiña, C. (2011). A Conceptual Model for Cultural Heritage Definition and Motivation. In *CAA'11: 39th Annual Conference on Computer Applications and Quantitative Methods in Archaeology*. Beijing, China.

Heer, J. A., Maneesh, A. (2006). Software Design Patterns for Information Visualization. In *IEEE Trans. Vis. Comput. Graph.* 12(5): 853-860.

Hobbs, Jerry R. (1985). On the coherence and structure of discourse. In *Technical Report*, Center for the Study of Language and Information (CSLI), Stanford, CA.

Huang, Z., Chen, H. et al. (2006). Expertise visualization: An implementation and study based on cognitive fit theory. In *Decision Support Systems* 42(3): 1539-1557.

Lacson, R.C.; Barzilay, R.; Long, W.J. (2006). Automatic analysis of medical dialogue in the home hemodialysisi domain: Structure induction and summarization. In *Journal of Biomedical Informatics* 39(5): 541-555.

Martín-Rodilla, P. (2012). The role of software in cultural heritage issues: Types, user needs and design guidelines based on principles of interaction. In *RCIS'12: Sixth International Conference on Research Challenges in Information Science*, Valencia, Spain.

McGrath, L.; Kuteeva, M. (2012). Stance and engagement in pure mathematics research articles. Linking discourse features to disciplinary practices. In *English for Specific Purposes* 31(3): 161-173.

Moens, M.F.; Boiy, E.; Mochales Palau, R.; Reed, C. (2007). Automatic Detection of Arguments in Legal Texts. In *ICAIL '07 Proceedings of the 11th International Conference on Artificial Intelligence and Law*. California, USA.

Morato, J; Llorens, J.; Genova, G.; Moreiro, J.A. (2003). Experiments in Discourse Analysis Impact on Information Classification and Retrieval Algorithms. In *Information Processing and Management* 39(6): 825-851.

North, C.; Schneiderman, B. (2000). Snap-Together Visualization: A User Interface for Coordinating Visualizations via Relational Schemata. In *AVI '00 Proceedings of the working conference on Advanced Visual Interfaces*. Palermo, Italy.

Prieto-Martínez, M. P. (1998). Forma, estilo y contexto en la cultura material de la Edad del Bronce gallega: cerámica campaniforme y cerámica no decorada, PhD Thesis. University of Santiago de Compostela.

Stasko, J., Zang, E. (2000). Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *InfoVis'00: IEEE Symposium on Information Visualization*, Salt Lake City, UT.

Tufte, E. R. (1990).*Envisioning Information*. Cheshire, CT: Graphics Press.

Van der Vlist, B.; Bram, J.J.; Niezen, Gerrit; Hu, Jun et al. (2011). Interaction Primitives: Describing Interaction Capabilities of Smart Objects in Ubiquitous Computing Environments. In *IEEE AFRICON 2011*, Livingstone, Zambia, IEEE.

Van Wijk, J. J.; Van de Wetering, H. (1999). Cushion treemaps: visualization of hierarchical information. In *InfoVis'99: IEEE Symposium on Information Visualization*, San Francisco, CA.

Vote, E., Acevedo Feliz, D., Laidlaw, D H.; Sharp Joukowsky, M. (2002). Discovering Petra: Archaeological Analysis in Vr. In *IEEE Computer Graphics and Applications* 22(5): 38-50.

Yi, J.S., Kang Y., Stasko, J.T., Jacko, J.A. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. In *IEEE Transactions on Visualization and Computer Graphics* 13(6): 1224-1231.

Zhou X., M.; Feiner, S. K. (1998). Visual task characterization for automated visual discourse synthesis. In *CHI '98 Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM Press/Addison-Wesley. Los Angeles, CA.