Polowinski, Jan, und Martin Voigt. „VISO: A Shared, Formal Knowledge Base as a Foundation for Semi-automatic InfoVis Systems“. In CHI ’13 Extended Abstracts on Human Factors in Computing Systems, 1791–1796. CHI EA ’13. New York, NY, USA: ACM, 2013. http://doi.acm.org/10.1145/2468356.2468677

VISO: A Shared, Formal Knowledge Base as a Foundation for Semi-automatic InfoVis Systems

Jan Polowinski
Technische Universität Dresden
01062 Dresden, Germany
jan.polowinski@tu-dresden.de

Martin Voigt
Technische Universität Dresden
01062 Dresden, Germany
martin.voigt@tu-dresden.de

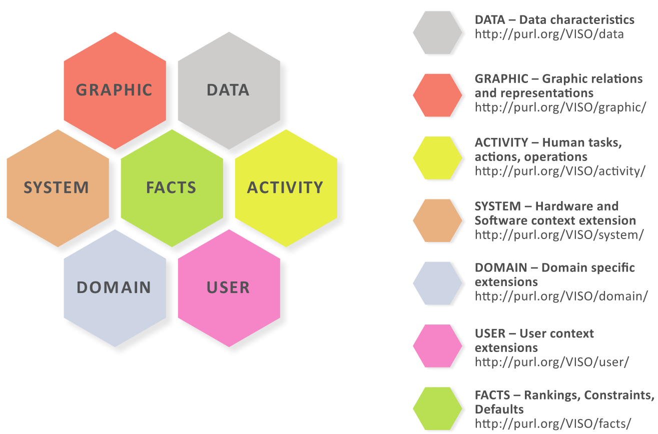

Figure 1: The Visualization Ontology (VISO) is a composite of seven modules, each focusing on a different field of visualization

CHI 2013 Extended Abstracts, April 27–May 2, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract

Interactive visual analytic systems can help to solve the problem of identifying relevant information in the growing amount of data. For guiding the user through visualization tasks, these semi-automatic systems need to store and use knowledge of this interdisciplinary domain. Unfortunately, visualisation knowledge stored in one system cannot easily be reused in another due to a lack of shared formal models. In order to approach this problem, we introduce a visualization ontology (VISO) that formally models visualization-specific concepts and facts. Furthermore, we give first examples of the ontology’s use within two systems and highlight how the community can get involved in extending and improving it.

Author Keywords

Ontology; knowledge; information visualization; interoperability

ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous.

Introduction

Due to the tremendous growth of data in recent years, it has become more and more challenging to identify relevant information in data. Interactive visual analytic systems can help to tackle this problem. Unfortunately, these systems suffer from a lack of interoperability and the knowledge used for guiding the user through visualization tasks cannot be shared and reused between these systems. Furthermore, various visualization approaches use different names for the same concepts, causing confusion also on the side of the user. As an example take the term *Graphical Technique* – two users may have two different concepts in mind when using this term. Similarly, two different users or systems may use varying terms for the same concept (e.g., *Visual Attribute* vs. *Visual Variable*).

We argue that a collaboratively developed ontology which formalizes common concepts, relations, and facts from the broad corpus of visualization literature and which is actively discussed in the community will be a solid foundation for upcoming (semi-)automatic visualization systems. Important advantages are (1) technical interoperability, supporting reuse of visualization knowledge, (2) the support of a common understanding between all interdisciplinary stakeholders in the visualization process, and (3) the ability to derive new knowledge from existing facts. We expect that this can lead to more accepted and "intelligent" applications.

In this paper, we give a brief overview of challenges for creating such a knowledge-base, present the current version of our visualization ontology (VISO), highlight how the community can get involved, and introduce how it is used as a foundation of two visualization systems in the domain of the Semantic Web.

A Common Visualization Ontology

Analysis and Preparation

In preparation of our knowledge base, we compared a broad corpus of articles from the field of graphics theory and information visualization, especially those already suggesting (informal) classifications and taxonomies (53 sources). Although, the need for formal visualization ontologies has been identified and first approaches were discussed [3], these are too abstract for our purposes or lack accessibility and reusability.

In a second step, before actually starting the modeling, we manually mapped the concepts used in literature to find synonyms, homonyms and term overlapping. A detailed description of this analysis is described in [7].

VISO in a nutshell

The Visualization Ontology (VISO) formalizes knowledge from the domain of visualization, in order to make it usable by machines and allow for exchange between tools and users. Machine-readability and interoperability is achieved using well-established Semantic Web standards such as RDF(S) and OWL [2].

VISO is modularized into seven parts (Fig. 1). The most important modules are GRAPHIC – formalizing terms such as *Graphic attribute* and *Graphic representation*, DATA – allowing to characterize data variables and structures, and ACTIVITY – being concerned with the human aspects of visualization, i.e. tasks, actions and operations. SYSTEM, USER and DOMAIN allow for describing the visualization context and domain-specific facts. The FACTS module formalizes constraints and rankings, e. g., of graphic relations, that have been described in literature and makes this knowledge available to tools in a standardized, interoperable way.

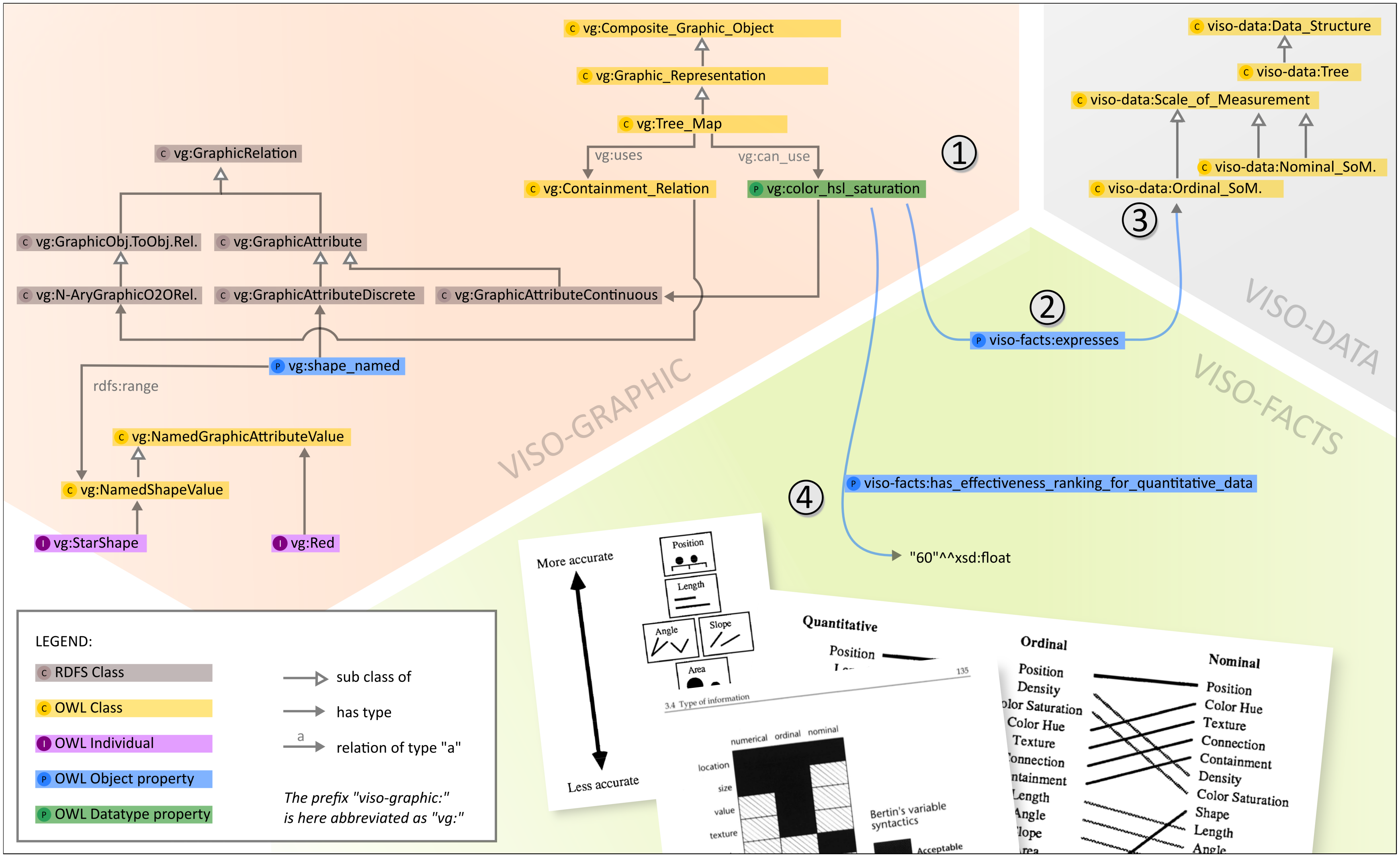

Figure 2: This diagram introduces some of the terms defined by the three most important modules (GRAPHIC/DATA/FACTS) and shows how they are connected using a concrete example: For the graphic attribute *Saturation* (viso-graphic:color_hsl_saturation) ①, it is stated in the FACTS module ② that saturation can express data with an *Ordinal scale of measurement* ③. Furthermore, saturation is assigned an effectiveness value for quantitative data of *"60"* (on an ordinal scale) ④.
OWL and RDFS Classes are comparable to (yet not the same as) classes in object-oriented programming, while OWL individuals can roughly be thought of as instances. OWL Object and Datatype properties model relations and attributes. The images illustrating further rankings defined in the FACTS module are taken from [4, 5].

In the remainder of this paper, we use the modules GRAPHIC, DATA and FACTS to demonstrate how VISO formally stores visualization-specific knowledge. Fig. 2 shows a subset of classes, instances and properties defined in these three modules and illustrates how they are connected. The GRAPHIC module is shown in most detail to provide an idea of how resources in a module are linked with each other. We show the sub-class hierarchy of the class *Graphic relation* which differentiates into *Graphic attribute* and *Graphic Object-to-Object relation* (Fig. 3).

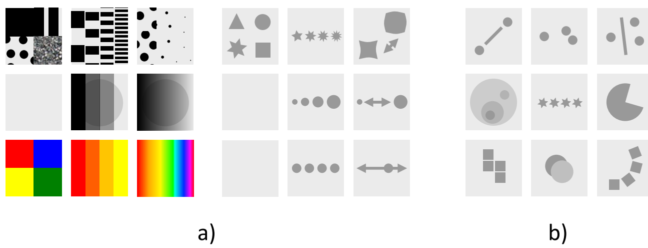

Figure 3: An excerpt of graphic relations defined in the GRAPHIC module of VISO. These include both *graphic attributes* (a) such as *color, texture* and *shape* and relativ graphic relations between graphic objects, we refer to as *graphic-object-to-object-relations* (b) such as *linked to*.

Also, concrete relations are modeled such as *Containment* as well as discrete and continuous attributes such as *Shape (named)* or *Saturation* (in the HSL color model). The DATA module defines terms like *Scale of Measurement*, *Data Structure* and their sub-classes. Finally, the FACTS module provides properties to relate terms from other modules according to facts that we found in the literature, e. g., in the rankings of effectiveness and efficiency of graphic relations by Mackinlay [5]. Due to space limitations, only a very small part of the DATA and FACTS module can be shown in Fig. 2. As a concrete example of how the three modules are related, look at the graphic attribute *viso-graphic:color_hsl_saturation* ①. For saturation it is stated in the FACTS module ② that it can *express* data with an *Ordinal Scale of Measurement* ③.

Documentation and References to Literature

A detailed documentation of each ontology module is automatically generated using the LODE ontology-documentation tool[1] and includes a description of each ontology term, its related terms, and depictions.

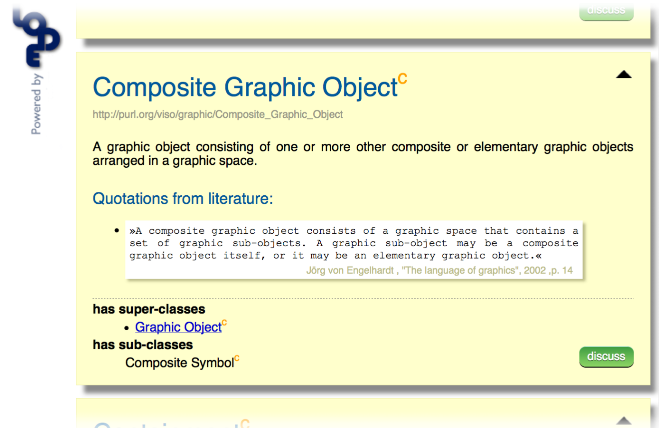

Figure 4: Excerpt of the VISO documentation showing the term *Composite Graphic Object*. Quotations from literature are given as annotations and a link is offered to discuss the term in the forum.

[1] http://lode.sourceforge.net/

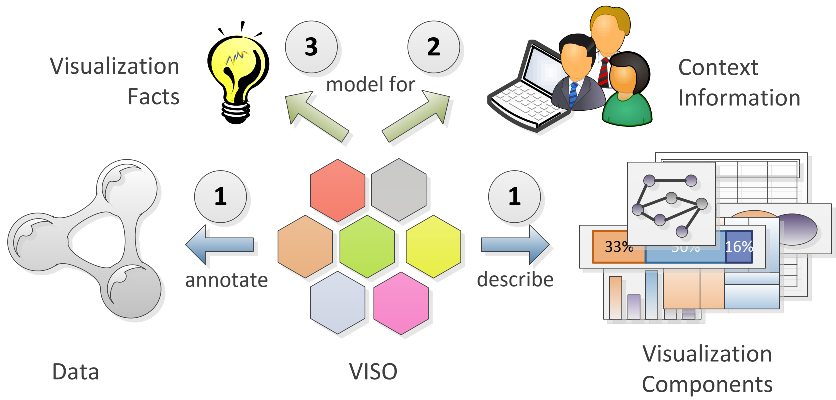

Figure 5: Overview of the usage of VISO within VizBoard

We extended LODE to show examples as well as quotations and literature references for each resource (Fig. 4). Since quotations, references, and authors are modelled in RDF as well, they can be conveniently queried using the RDF query language SPARQL [1] , e. g., to find out what are the terms a specific author has influenced.

Get Involved
A beta version of VISO can be downloaded from http://purl.org/viso/. Pointing your web-browser to the same URL will show you the documentation instead.

Contributing to the "shared" aspect of our knowledge base, we planned from the beginning to allow for participation of the visualization community. We started by providing a question & answers platform which is integrated with the ontology documentation. Each term can be discussed by the community, existing definitions can be ranked and new interpretations can be proposed.

VISO as a Foundation for InfoVis Systems

In the following, we introduce two exemplary approaches relying on VISO.

VizBoard – a generic InfoVis workbench for semantic data – tackles (amongst others) the problem of recommending proper graphic representations within a specific visualization context [6]. Therefore, the VISO vocabulary is employed in various ways (Fig. 5). First, to describe visualization components and to annotate the data according to their characteristics. Second, VISO is the foundation to store context information about the user, his task and device. Third, the knowledge formalized by means of the *FACTS* module allows for identifying and ranking suitable visualization widgets.

A second example for the usage of VISO is the RDF Visualization Language (RVL)[2]. RVL allows for defining declarative mappings between domain properties – described in the Semantic Web languages RDF(S) and OWL – and properties of the VISO GRAPHIC module such as "color_hsl_lightness" or "Containment_Relation". That means, VISO is used to model the graphic relations that form the target of a mapping definition. Each mapping gets its own URI which supports sharing, reuse and composition of mappings.

VISO could also be employed for other purposes in future: First, it may be used to classify visualizations, thereby supporting the search for visualizations and papers as well as the discovery of "under-researched" areas. Second, it may help to consolidate vocabulary used in the field of visualization by clarifying synonyms, homonyms, and term overlap.

[2] http://purl.org/rvl/ ; a technical report on RVL will be published in early 2013

Conclusion and Further Work

We presented the current state of the VISO ontology, a shared, formal knowledge-base on visualization, and showed what it can already be used for. When building your own interactive InfoVis system you have different options to employ VISO. A simple but effective example is linking to VISO resources by their URI, e. g., http://purl.org/viso/graphic/Graphic_Representation, which provides a label and description in multiple languages as well as the instances and specialisations for the selected concept. As a developer, you can reuse this knowledge – which may help users to understand visualisation terms – instead of providing it on your own. Beyond this simple usage, you could also benefit from the rankings offered by the VISO/FACTS module to suggest appropriate graphic relations for your data.

As ontologies represent shared knowledge and are always a work-in-progress, we encourage other researchers from the field of visualization and human computer interaction in general to discuss the terms we chose for the initial version of the ontology, in order to yield a both broadly accepted and logically consistent knowledge base. You are welcome to contribute to the VISO development process by criticizing, suggesting new extensions, or joining the developers.

Acknowledgements

This research has been co-funded by the European Social Fond / Free State of Saxony, contract no. 80937064, 1330674013 (eScience – network), and 99457/2677.

References

[1] SPARQL 1.1 query language. http://www.w3.org/TR/sparql11-query/, Jan. 2012.

[2] Bechhofer, S., Van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., and Stein, L. A. OWL web ontology language reference. *W3C recommendation 10* (2004).

[3] Brodlie, K. W., Duce, D. A., Duke, D. J., et al. Visualization ontologies: Report of a workshop held at the national e-science centre. Tech. rep., e-Science Institute, Edinburgh, Scotland, 2004.

[4] Engelhardt, von, J. *The Language of Graphics*. PhD thesis, Institute for Logic, Language & Computation, University of Amsterdam., 2002. ISBN 90-5776-089-4.

[5] Mackinlay, J. Automating the design of graphical presentations of relational information. *ACM Trans. Graph. 5*, 2 (1986), 110–141.

[6] Voigt, M., Pietschmann, S., Grammel, L., and Meißner, K. Context-aware recommendation of visualization components. In *Proc. of the 4th International Conference on Information, Process, and Knowledge Management (eKNOW 2012)*, XPS (Feb. 2012).

[7] Voigt, M., and Polowinski, J. Towards a unifying visualization ontology. Technical Report TUD-FI11-01, TU Dresden, Institut für Software und Multimediatechnik, Dresden, Mar. 2011. ISSN: 1430-211X.