Semantic Visualisation of Heterogenous Knowledge Sources

Hans-Georg Fill
Institute of Knowledge and Business Engineering, University of Vienna
hans-georg.fill@dke.univie.ac.at

Abstract: This paper presents a novel approach in the area of knowledge visualisation for integrating heterogenous knowledge sources. For this purpose a semantic visualisation framework is presented that provides conceptual foundations for the creation of semantically described visualisations. The concepts are presented formally and have been used to develop a first prototype implementation for giving a proof of concept.

1 Introduction

The use of modelling methods for representing knowledge is common in many business areas today: Examples can be found amongst others in business process management [KNS92, KJS96], strategic management (e.g. for the representation of strategic goals and their relationships [LKK02]), or e-learning management (e.g. for realising the vision of the learning organisation [BS03]). The central concept is to provide a standardised set of elements that are either used in their original form or act as a template for a user to depict his or her personal view of a domain. By explicating this knowledge it can be shared with others for communication, integration, and enhancement purposes.

Although there exist various approaches for representing such models to a user, including textual descriptions, formal or semi-formal methods (e.g. the wide range of methods from logics and artificial intelligence) or purely structural descriptions (e.g. from the field of databases) the availability of visualisations of such models has contributed much to their ease of use and thus to their wide acceptance. Visualisations can therefore be seen as an integral part in modelling and exchanging knowledge.

But unlike other approaches for representing knowledge where the syntax, semantics and pragmatics for methods are thoroughly researched and often internationally aligned in the form of standards (as e.g. XML as a basis for representing syntax and RDF, RDFS, OWL for semantics) similar efforts for visualisations of knowledge are currently not that apparent. To gain insight into the fundamentals of visual representations for such models one has to revert to a number of different research fields (as discussed in section 2) and bring together the existing insights. To overcome this deficiency the following remarks aim to provide a starting point for a coherent view on the process of visualisation from a knowledge oriented view and illustrate an approach how the linkage of visual representations to semantic definitions can be realised.

The intended goal is to establish semantically described visualisations that can be automatically linked to domain models. The anticipated advantage would be to allow for an easier understanding of the concepts and relationships expressed in the models and thereby ease the process of integrating different knowledge sources with visual means.

2 Related Work

As mentioned in the introduction the basic principles for the approach outlined here have to revert to a number of different scientific fields thereby making it highly interdisciplinary: In the area of knowledge engineering a well known approach for visually representing knowledge is to revert to graphs (e.g. conceptual graphs cf. [Sow00], entity-relationship diagrams or a number of approaches that recently came up for visualising ontologies cf. [FSH03, LS03]). Although these approaches provide in some cases highly complex 2D and 3D visualisations of semantic relationships, from a pragmatical point of view they are still very far away from what a user that is not very familiar with the underlying formalisms would understand. A simple example shall illustrate this: Consider an ontology that describes the concepts of temperatures (e.g. in Fahrenheit, Celsius and Kelvin) and relates them to the concept of a production machine to evaluate whether its operation temperature is right. If these relationships are expressed as conceptual graphs for example they are hard to understand but if the concept of temperature could be automatically linked to a thermometer visualisation that displays the values (as a typical user would expect) the relationships would not need much explanation.

To integrate such aspects in visualisations of ontologies for example it would be necessary to know the semantics of the symbols involved and to be able to express them in a computer processable, i.e. formal, way as well. The science that studies the meanings of signs in general [Kaz95] and its cultural embedding [Sch97] is semiotics. However this science is not particularly related to computer applications. Semiotics from an IT point of view are partly covered by field of human-computer interaction and visual language theory [BHLT04]. Visual language theory addresses the needs for an exact, formal specification of visualisations [NH98] and is strongly related to the concept of ontological visualisation patterns as discussed in section 4. Although it forms the basis of the knowledge oriented view of visualisation presented in this paper so far no internationally spread standard for semantically exact specifications for the visual representations of visual languages has yet been proposed.

The field of information visualisation which is defined as "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [CMS99](p. 7) uses partly principles from visual language theory and semiotics as well as from computer graphics and creates new methods for representing information in a visual form [War00]. However very often the focus here is on large, high-dimensional data sets [Grö01]. This is especially true for the closely related area of scientific visualisation where scientific data that is generated by experiments needs to be understood which is often not possible or feasible with traditional data analysis. Information visualisation is also concerned about the automation of the visualisation process as e.g. described among others by [RM90] and [Mac86] who provide results on the coupling of qualitative and quantitative data to visual representations. Despite their high international reputation the deficiency of these fields for the intended purposes in the visualisation of knowledge models lies in the stressing of a basis of data resp. information which is traditionally not equal to knowledge [Nor02]. For the term knowledge as it shall be defined here the important characterisation is the human aspect - ”humanised information” as it has been coined by D. Karagiannis - i.e. the human interpretation of information and the resulting dependencies e.g. from a semiotic and cultural point of view. So it’s not only the syntax and semantics of the data that is relevant but also the semantics and implications of the chosen visual representation.

The comparatively young field of knowledge visualisation [Bur04, Epp04] takes these aspects into account to some extent and stresses the communicative aspect of visualisations: It is defined as the ”use of visual representations to improve the transfer of knowledge between at least two persons or group of persons.” [Bur04].

The area of computer graphics [HB97] supplies the technical and mathematical basics for a realisation of visualisations with information technology. Direct achievements of this field have not only had a major impact on the technical, natural and economic sciences making use of visualisation but has also substantially influenced the entertainment industry that can today be regarded as the driving force for the technical and economical advancement in graphics hard- and software [Rhy02].

Further extensions of the approach presented in this paper can be made by referring to the fields of human-computer interaction, psychology (especially in regard to visual perception and pedagogical issues) and multimedia technologies which have not been particularly considered in first place.

3 Methods for Specifying Domains

The basis for representing knowledge with the support of information technology is the availability of formal or at least semi-formal methods for defining the elements and structure of a domain. For this purpose the concept of ontologies is currently widely discussed in academia and starts to gain ground in industry. Ontologies can be generally defined as ”shared sets of explicitly defined terminology” [NFF$^+$91](p.38) and the sub-type of domain ontologies as ”an integral and coherent conceptualisation of a domain” [BA97](p.378). Ontologies extend the idea of exact vocabularies and classifications of terms of a domain by interlinking the described concepts and are built upon formally exact specifications. In the context of the Semantic Web, ontologies play an essential role for realising machine processable semantics and derived applications such as intelligent agents [BLHL01].

A similar approach that is based on the same concept of specifying a domain is to use meta models [KK02]. The difference to ontologies lies in their degree of abstraction from the real world: Whereas the quality of ontologies increases the more concepts are referenced and related to each other[1], meta models usually abstract from the real world to a greater extent and aim to present a manageable view of the world to a user.

[1]This view assumes a ’network effect’ for ontologies, i.e. the more parts of the world are described by ontologies the better and more detailed conclusions based on the ontology can be made.

Both approaches ontologies and meta models have in common that their basic components are atomic concepts or unary predicates and atomic roles or binary predicates that relate the atomic concepts to each other. This structure results from the fundamentals of description logic [BHS04](p. 5) and directly leads to the above mentioned graph representations where nodes stand for atomic concepts and edges for atomic roles. In the following it is assumed that there exists an ontology or a meta model of a domain of interest that a visualisation shall be based on.

4 A Framework for Semantic Visualisation

To provide answers to the open questions given in the introduction several paths can be chosen to achieve successful results. The use of information technology is one path in this context but not necessarily the only successful one. Therefore the aim of the subsequently presented framework (see figure 1) is to conceptually integrate IT based views as well as non IT views to visualisation.

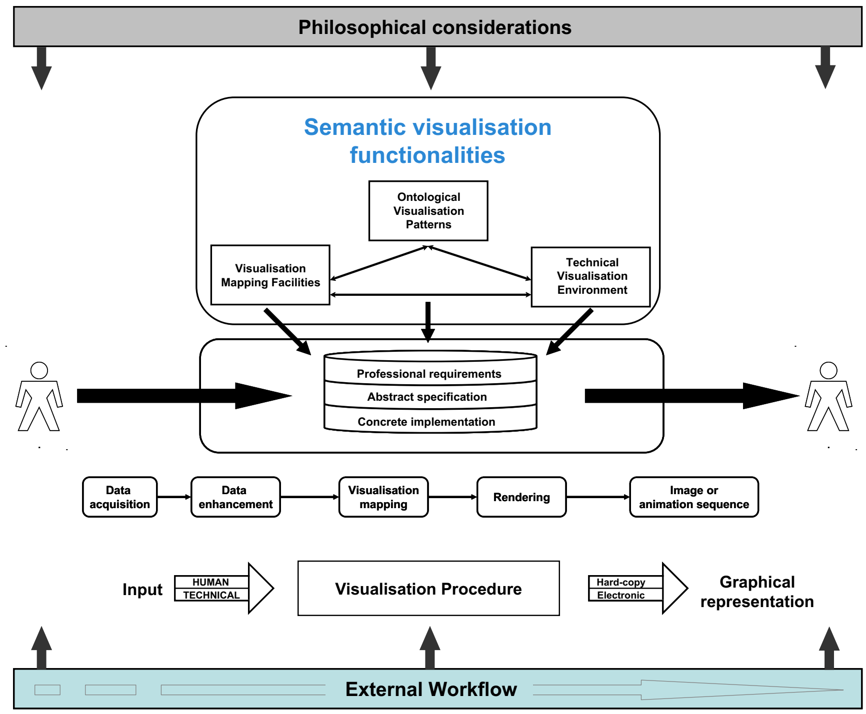

Figure 1: Semantic Visualisation Framework

The process of the generation of a visualisation is estimated to be influenced on the bottom by an external workflow, i.e. a sequence of activities that encompasses the visualisation task and philosophical considerations at the top, that incorporate scientific concepts and thoughts related to visualisation (concretely the philosophical considerations by Euclid for defining geometric figures on the basis of the uniform entity of the point have influenced the conceptions of visual objects as laid out below).

The output of the process is a graphical representation that may be available as hard-copy or in an electronic format. The output is generated by the abstract concept of a visualisation procedure (which will be detailed in the following) that again is supplied with either a human or a technical input (i.e. abstract data or human knowledge as for example expressed by arranging visual elements in a modelling tool). This separation of the input therefore takes into account both the data based approach of information visualisation and the knowledge oriented approach by knowledge visualisation.

When using information technology for creating visualisations the exact steps for the creation of a visualisation can be described by the visualisation pipeline of computer graphics [Grö01] which is shown above the generic visualisation process in figure 1. It describes the transformation of data to visualisations by assigning the enhanced data to visual properties and then rendering them on computer graphics hardware so that an image or a sequence of images are shown. Thereby the visualisation mapping is the central part that determines the actual appearance of the representation.

To realise the vision of semantically defined visualisations the visualisation pipeline as a concept (not yet as a technical instantiation) has to be extended. The first extension concerns the integration of the aspect of human knowledge that is seen as the overall input and output of the visualisation process. The reason is that (at least until now) visualisations are only for human purposes - so the resulting characteristics have to be kept in mind (symbolised by the human symbols at the far left and right).

The remaining extensions are comprised of two main parts that are shown above the pipeline in figure 1: Semantic visualisation functionalities and a new visualisation procedure. The new visualisation procedure takes into account the existence of abstract specifications for possible visualisations (visualisation templates). Their properties are determined by professional requirements of a particular application area and influenced by the context of the concrete visualisation task in the form of the external workflow. When the abstract specification is instantiated, a concrete implementation is generated that is available to the user of the visualisation. Thus the visualisation procedure is not narrowed to the technical realisation but is raised to a conceptual level that links the procedure to the professional requirements and explicitely defines abstract specifications that allow for a re-use or if formally described for an automatic generation of the visualisation.

These three elements are supported by three inter-connected semantic visualisation functionalities:

- Ontological Visualisation Patterns
- Visualisation Mapping Facilities
- Technical Visualisation Environment

The term *Ontological Visualisation Patterns* constitutes a new way of defining visualisations. Ontological Visualisation patterns contain several features that have not been conceptualised before: They are the main factor for fulfilling the requirements of a service oriented visualisation architecture where small (software) components can be accessed for providing a specific visualiation functionality. Furthermore they contain abstract specifications of visualisations that are semantically specified in a standardised form and can therefore be directly used by intelligent agents.

For illustrating their relation to concrete graphics the subsequent illustration has been created in analogy to the well known semantic web stack (see figure 2). On its bottom layer it contains the core elements of any visualisation: *Points* as the smallest visual entity (influenced by philosophical considerations and the necessity for the use of points due to rasterised displays in computer graphics) and *Graphical Primitives* that are composed of a number of points. On the next layer reside *Graphical Representations* in the form of vector graphics that are composed of graphical primitives[2]. To allow for a dynamic change of the graphical representations the *Visual objects* layer adds variable properties and control structures for these properties to the graphical representations. Thereby a representation can be changed according to a variably assigned value. An example would be a rectangle whose length is specified by a variable and that can thus be modified arbitrarily. On the *Ontological Visualisations Patterns* layer one or more visual objects are combined and positioned according to a layout procedure. The variables of the visual objects are either directly exposed as variables of the pattern or via a transformation function. The ontological visualisation pattern therefore also has to contain control structures and variables, but in contrast to the visual objects the variables as well as the pattern itself are enriched with a semantic specification. On the basis of this semantic specification that describes the nature and properties of a visualisation a mapping to an ontology can be established that is the basis for possible visual inferencing mechanisms for *Logic*, *Proof* and *Trust*.

A formalisation for these concepts and the linking to a semantic description could be as follows. A Point P has coordinate values in $\mathbb{R}^2$ (1), and a variable V in $\mathbb{R}$ (2). Primitives *PRIM* can either be described by points (3) or in visual objects by a combination of points and variables *VO* (4). This abstract specification can then be instantiated (5).

A simplified ontology O consisting of classes C, object properties OP, datatype properties DP and values V (7)[3] can also be instantiated and then linked in the simplest case to the visual object as shown in (13,14), thereby establishing a semantic relationship with a class of the ontology and a value.

$$
\begin{aligned}
P &= (x, y \,|x, y \in \mathbb{R}) && (1)\\
V &= (z \,|z \in \mathbb{R}) && (2)\\
\mathit{PRIM} &= \{P_1, P_2, \ldots, P_n\} && (3)\\
\mathit{VO} &= \{E_1, E_2, \ldots, E_n \,|E \in P \vee V\} && (4)\\
\mathit{VO}_1 &:: \mathit{VO} && (5)
\end{aligned}
$$

[2]Vector graphics are choosen to assure the lossless manipulation of the graphics.
[3]This would correspond to a simplified conception of the Web Ontology Language OWL.

$$
\begin{aligned}
VO_1 &= \{P_1, P_2, V_1\} && (6)\\
O &= \{C, OP, DP, VAL\} && (7)\\
C &= \{C_1, C_2\} && (8)\\
OP &= \{OP_1\} && (9)\\
DP &= \{DP_1\} && (10)\\
OP_1 &= \{C_1, C_2\} && (11)\\
DP_1 &= \{C_2, VAL_1\} && (12)\\
C_2 &\leftrightarrow VO_1 && (13)\\
VAL_1 &\leftrightarrow V_1 && (14)
\end{aligned}
$$

To complete the concept of semantic visualisation functionalities the *Visualisation Mapping Facilities* are required to be responsible for linking the ontological visualisation patterns to the external environment. This could for example be done by intelligent agents that can match the semantic descriptions of the pattern with a concrete ontology.

With the functionalities of the *Technical Visualisation Environment* the above described functionalities can be implemented. Again, seen from the conception, this part does not need to be equivalent to computer science technologies. It shall be stressed that results can also be achieved manually (then the framework would be instantiated e.g. by the use of a number of coloured pens and a sheet of paper, the ontological visualisation patterns would then be a kind of stencil where the shapes can be dynamically adapted).

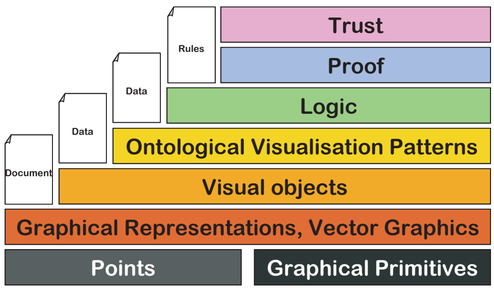

Figure 2: Semantic Visualisation Stack

5 Possible usage scenarios

The coherent view on the process of visualisation and its influence factors as it is described by the approach of the Semantic Visualisation Framework allows for a meta-level description and positioning of different visualisation approaches. By instantiating the Semantic Visualisation Stack the provision of semantic visualisation services becomes feasible that embody platform independent, loosely coupled, non-monolithic, machine consumable software components.

To illustrate these aspects two usage scenarios shall be described in the following: In several approaches for performance management, visualisations are used to indicate the achievement of goals. Very often graphical representations such as thermometers (as a metaphor for the right and deviating values) or tachometers (with red, yellow and green areas symbolising not achieved, partly achieved and fully achieved target values) are integrated in these 'management cockpits'(e.g. [Geo00]). The creation of new metaphoric representations for such cockpits in traditional approaches necessitates the programming of software components which requires programming experts as well as the adaption to a specific tool or technology. By using visual objects that incorporate the needed visualisation functionalities (e.g. for altering the colour of the object or the indicated value in the tachometer) and linking the objects to an ontology for performance management that conceptualises the range of possible target values and their meaning, the visualisations can be used by different tools and technologies. Through offering visual objects via a semantic visualisation web service these visualisation functionalities could also be obtained automatically from the internet.

For the second example consider a new visualisation method for the representation of business processes that allows for the integration of relative costs of elements into the visual representations, e.g. by gradually colouring the process activities from black to light red according to the ranking of their costs in the entire model (an example for such a visualisation method is described in detail in [FH06]). Traditional modelling approaches for business process management such as ADONIS BPMS [FH06] or event driven process chains [KNS92] already consider the cost attributes necessary for the visualisation in their meta-models and thus fulfill the requirements for the syntax and semantics of a modelling language but might fall short of providing visualisation methods for integrating these aspects into the graphical notation.

One possible solution would therefore be to change the notation of the modelling method (e.g. in the case of ADONIS as described in [Fil04]). However, this approach usually requires programming or at least customisation effort. By using the approach of Semantic Visualisation a visualisation can be specified separately from the meta-model of a modelling language and then linked to the meta-model on demand. In this scenario the formal specification of visualisations provides the necessary precision for unambiguous and interchangeable definitions of the resulting graphical representations (which is not possible with currently available notation definitions such as the Business Process Modelling Notation [Whi04]) as well as the basis for semantically linking the cost attribute of the meta-model/ontology to the colour variable of the visual object that is used to represent an activity. Furthermore, due to meta-level and technology independent approach a re-use of visualisation objects is facilitated.

6 Prototypical Implementation of Visual Objects

The Semantic Visualisation Framework in the previous section has been instantiated partially by the author using information technology as a technical visualisation environment. A first approach for ontological visualisation patterns based on visual objects has been realised by the development of a visual editor that is capable to enrich graphical representations with variables and control structures (which would not currently be possible using technology independent graphics standards such as SVG). The screenshot in figure 3 shows the Java-based visual object editor and a visual object on the right side. Currently visual objects can be created using a number of primitives that can then be enriched with variables to allow for their dynamic modification. The semantic description of the visual objects has then (so far) to be performed manually (e.g. by adding RDF statements to the variables).

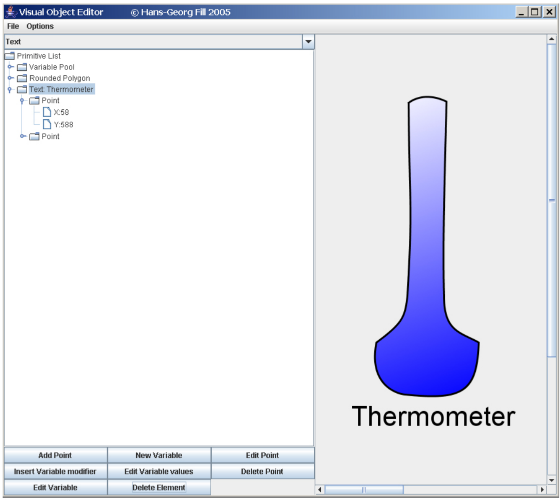

Figure 3: Screenshot of the visual object editor

7 Conclusion and further work

In this paper a new approach for specifying visualisations semantically has been elaborated by presenting a framework for semantic visualisation and illustrating the main functionalities by a prototype implementation. It is estimated that based on these results a first step towards a visual semantic integration of knowledge sources has been made that can be the basis for a new, visual driven direction in the field of knowledge modelling.

Future work will include the extension of the prototype implementation as well as an instanciation of the framework building upon extensions of existing graphic standards.

References

[BA97] P. Borst and H. Akkermans. Engineering ontologies. *International Journal of Human-Computer Studies*, 46:365–406, 1997.

[BHLT04] P. Bottoni, C. Hundhausen, S. Levialdi, and G. Tortora, editors. *Proceedings of the 2004 IEEE Symposium on Visual Languages and Human Centric Computing*. IEEE Computer Society, 2004.

[BHS04] F. Baader, I. Horrocks, and U. Sattler. Description Logics. In S. Staab and R. Studer, editors, *Handbook on Ontologies*. Springer, Berlin - Heidelberg, 2004.

[BLHL01] T. Berners-Lee, J. Hendler, and O. Lassila. The Semantic Web - A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, May 2001, 2001.

[BS03] J. Bajnai and C. Steinberger. eduBITE Educating Business and Information Technologies. In *E-Learn 2003, World Conference on E-Learning in Corporate Government, Healthcare & Higher Education*, pages 863–866. AACE, 2003.

[Bur04] R. A. Burkhard. Learning from Architects: The Difference between Knowledge Visualization and Information Visualization. In *Draft for 8th International Conference on Information Visualization (IV04)*, London, 2004.

[CMS99] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in Information Visualization - Using Vision to Think*. Morgan Kaufmann, San Francisco, CA, 1999.

[Epp04] M. J. Eppler. Facilitating Knowledge Communication through Joint Interactive Visualization. In *I-KNOW '04*, Graz, Austria, 2004.

[FH06] H.-G. Fill and P. Höfferer. Visual Enhancements of Enterprise Models. In *accepted at: Multikonferenz Wirtschaftsinformatik 2006*, Passau, Germany, 2006.

[Fil04] H.-G. Fill. UML Statechart Diagrams on the ADONIS Metamodeling Platform. In *International Workshop on Graph-Based Tools (GraBaTs 2004)*, pages 27–36, Rome, Italy, 2004. Electronic Notes in Theoretical Computer Science.

[FSH03] C. Fluit, M. Sabou, and F. Van Harmelen. Ontology-based Information Visualization. In V. Geroimenko and C. Chen, editors, *Visualizing the Semantic Web - XML-based Internet and Information Visualization*, pages 36–48. Springer, London - Berlin - Heidelberg, 2003.

[Geo00] P. M. Georges. The Management Cockpit - the human interface for management software: reviewing 50 user sites over 10 years of experience. *Wirtschaftsinformatik*, 42(2):131–136, 2000.

[Grö01] E. Gröller. Insight into Data through Visualization. In P. Mutzel, M. Jünger, and S. Leipert, editors, *GD2001*. Springer, 2001.

[HB97] D. Hearn and P. M. Baker. *Computer graphics - C version*. Prentice Hall, Upper Saddle River, New Jersey, 1997.

[Kaz95] E. Kazmierczak. Universals in Visual Codification. In J. Bernard, editor, *Bildsprache - Visualisierung - Diagrammatik, Special Issue of* Semiotische Berichte*, Institute for Socio-Semiotic Studies Vienna*, volume 19, pages 131–152. Vienna, Austria, 1995.

[KJS96] Dimitris Karagiannis, Stefan Junginger, and Robert Strobl. Introduction to Business Process Management Systems Concepts. In Bernd Scholz-Reiter and Eberhard Stickel, editors, *Business Process Modelling*, pages 81–106. Springer, 1996.

[KK02] Dimitris Karagiannis and Harald Kühn. Metamodelling platforms. In Kurt Bauknecht, A Min Tjoa, and Gerald Quirchmayer, editors, *Proceedings of the third international conference ec-web 2002 - dexa 2002*. 2002.

[KNS92] Gerhard Keller, Markus Nüttgens, and August-Wilhelm Scheer. Semantische Prozeßmodellierung auf der Grundlage 'Ereignisgesteuerter Prozeßketten (EPK)'. In August-Wilhelm Scheer, editor, *Veröffentlichungen des Instituts für Wirtschaftsinformatik*. 1992.

[LKK02] C. Lichka, H. Kühn, and D. Karagiannis. ADOscore - IT gestützte Balanced Scorecard. *wisu-Das Wirtschaftsstudium*, (7):915–918, 2002.

[LS03] B. LeGrand and M. Soto. Topic Maps Visualization. In V. Geroimenko and C. Chen, editors, *Visualizing the Semantic Web - XML-based Internet and Information Visualization*, pages 49–62. Springer, London - Berlin - Heidelberg, 2003.

[Mac86] J. D. Mackinlay. Automating the Design of Graphical Presentations of Relational Information. *ACM Transactions on Graphics*, 5(2):110–141, 1986.

[NFF+91] R. Neches, R. Fikes, T. Finin, T. Gruber, R. Patil, T. Senator, and W. R. Swartout. Enabling Technology for Knowledge Sharing. *AI Magazine*, 12(3):36–56, 1991.

[NH98] N. H. Narayanan and R. Hübscher. Visual Language Theory: Towards a Human-Computer Interaction Perspective. In K. Marriott and B. Meyer, editors, *Visual language theory*. Springer, New York, 1998.

[Nor02] K. North. *Wissensorientierte Unternehmensführung: Wertschöpfung durch Wissen*. Gabler, Wiesbaden, 2002.

[Rhy02] T.-M. Rhynes. Computer games and scientific visualization. *Communications of the ACM*, 45(7):40–44, 2002.

[RM90] S. Roth and J. Mattis. Data characterization for intelligent graphics presentation. In *CHI'1990*. ACM, 1990.

[Sch97] A. Schelske. *Die kulturelle Bedeutung von Bildern: Soziologische und semiotische Überlegungen zur visuellen Kommunikation*. Deutscher Universitäts Verlag / Gabler, Vieweg, Westdeutscher Verlag, Wiesbaden, 1997.

[Sow00] J. F. Sowa. *Knowledge representation - logical, philosophical, and computational foundations*. Brooks/Cole, Pacific Grove, CA, USA, 2000.

[War00] C. Ware. *Information Visualization - Perception for design*. Morgan Kaufman, 2000.

[Whi04] S. A. White. Business Process Modeling Notation (BPMN) Version 1.0 - May 3, 2004. Technical report, BPMI.org, 2004. `http://www.bpmn.org/Documents/BPMN%20V1-0%20May%203%202004.pdf` accessed 06-02-2006.