# Modeling Interactive, 3-Dimensional Information Visualizations Supporting Information Seeking Behaviors

Gerald Jaeschke[1], Martin Leissler[2], and Matthias Hemmje[1]

[1] FernUniversität Hagen, Universitätsstr. 1,
58097 Hagen, Germany
{gerald.jaeschke, matthias.hemmje}@fernuni-hagen.de
[2] Brainmelt GmbH, Hugenottenallee 15,
63263 Neu Isenburg, Germany
martin@brainmelt.com

**Abstract.** Information visualization and knowledge visualization use comparable techniques and methods. Based on mapping rules, resource objects are translated into visual objects as meaningful representations, offering easy and comprehensive access. Whereas information visualization displays data objects and relations, knowledge visualization maps knowledge elements and ontologies. Bridging this gap must start at concept level. Our approach is to design a declarative language for describing and defining information visualization techniques. The information visualization modeling language (IVML) provides a means to formally represent, note, preserve, and communicate structure, appearance, behavior, and functionality of information visualization techniques and applications in a standardized way. The anticipated benefits comprise both application and theory.

## 1 Introduction

Knowledge visualization and information visualization have progressed independently of one another. But there are endeavors to bring these research fields together. Modern, computer-based mapping tools are capable of displaying both content and conceptual knowledge at the same time. In the predominant approach, conceptual maps structure the domain and serve as navigational tool that provides knowledge-based access to information. Links in the map attached to concepts reference underlying information and allow to immediately access that information. Information, also referred to as content knowledge, comprises, for example, personal notes, sketches, and example instances of concepts. With this linking feature, this kind of mapping tools could serve as an interface between knowledge visualization and information visualization when the content knowledge consists of abstract mass data. In such alliance, information visualization and knowledge visualization can collaborate side by side.

Information visualization and knowledge visualization can further mutually enrich each other beyond this scope. Today, both disciplines endue strong conceptual foundations. At the same time, information visualization and knowledge visualization

S.-O. Tergan and T. Keller (Eds.): Knowledge and Information Visualization, LNCS 3426, pp. 119 – 135, 2005.

employ comparable techniques and methods: Based on mapping rules, resource objects are translated into visual objects as meaningful representations, offering easy and comprehensive access to the subject matter presented. Whereas information visualization displays data objects and relations, knowledge visualization maps explicate knowledge representations, e.g. concepts.

Under these circumstances, knowledge visualization should be able to adopt achievements that emerged from information visualization research and vice versa. Mapping tools then could display conceptual knowledge and content knowledge within one and the same visual environment. Moreover, mapping tools could visualize knowledge applying (parts of) information visualization techniques and the other way round, provided that setting up a common basis is successful.

Bridging this gap must start at concept level. We tackle this challenge from the information visualization perspective. Our approach is to design a declarative language for describing and defining information visualization techniques. The information visualization modeling language (IVML) provides a means to formally represent, note, preserve, and communicate structure, appearance, behavior, and functionality of information visualization techniques and their applications in a standardized way.

The anticipated benefits comprise both application and theory. Standardized models allow for the specification and implementation of diverse interpreters serving various target platforms. Graphical user-interfaces deploying information visualization techniques can be described and dynamically generated on-the-fly, also by machines. More importantly, the underlying formal model underlying the dynamic generation also renders possible analysis and reasoning, in turn supporting the detection of (information) visualization design flaws.

Such a language needs to rest on solid foundations. The information visualization modeling language puts into practice a formal model that reflects the concepts and relationships of information visualization as it is understood today. To the best of our knowledge, no such integrated model exists. Research on information visualization has so far established an outline of the information visualization process and shed light on a broad range of detail aspects involved. However, there is no model in place that describes the nature of information visualization in a coherent, detailed, and well-defined way.

In order to mutually open-up the treasure chests of visualization techniques for knowledge and information, information visualization and knowledge visualization must base on a joint visualization model. Or at least, they must share a significant amount of visualization model. Integrating the principles of knowledge visualization techniques, such information and knowledge visualization modeling language (IKVML) could provide the means to represent both information and knowledge visualization techniques.

On our way towards the information visualization modeling language, first we survey and discuss extant models of which each covers selected facets of (information) visualization (section 2). The survey focuses on work that devised classification schemas. Our supposition that the presence of classifications indicate an elaborated level of formalization is the rationale behind this selection. Second, we provide an overview of the entire set of models under investigation and discuss the coverage of and the relationships between the models (section 3). Next, we present computational requirements as well as requirements imposed by the application the information visualization modeling language has to fulfill (section 4). We conclude by sketching application scenarios that illustrate the language's benefits within resource-based e-learning scenarios (section 5). Throughout this paper, we will refer to the visualization reference model in order to organize our investigations.

## 2 Information Visualization Models

"Classification lies at the heart of every scientific field." (Lohse, Biolsi, Walker & Rueter, 1994) In striving for a better understanding of information visualization, a variety of classification schemes have been proposed over the past years. Depending on provenance and intention, they shed light on the information visualization process, its application, or its utility. Information visualization techniques, applications, systems, and frameworks can be classified according to the data types they can display, user tasks they support, characteristics of visual representations they deploy as well as cognitive aspects of their visual appearance.

**Reference model for visualization.** Card, Mackinlay and Shneiderman (1999) introduced a reference model for information visualization (Fig. 1), which provides a high-level view on the (information) visualization process.

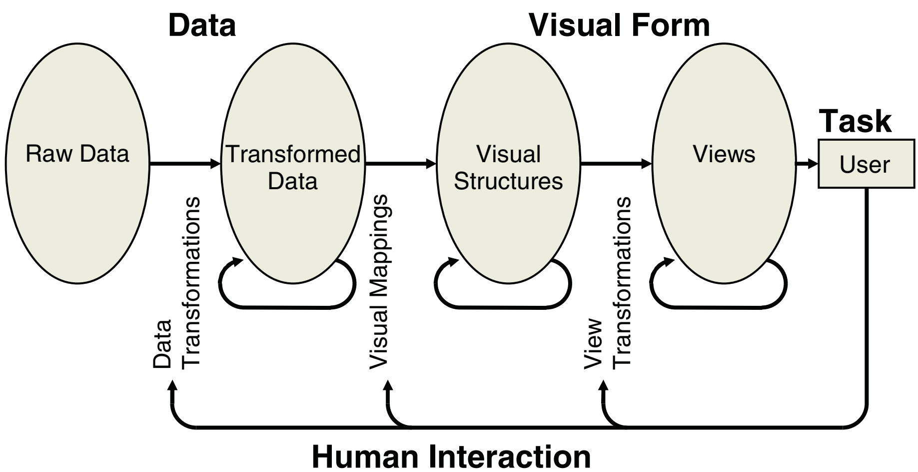


**Fig. 1.** Reference model for visualization

The model assumes a repository of raw data, which exist in a proprietary format, be it structured or unstructured. To get to a visualization of this data, data have to first undergo a set of transformations. Data transformations comprise filtering of raw data, computation of derived data as well as data normalization. These steps result in a set of transformed data in a unified structure. Visual transformations map the transformed data onto a corresponding visual structure. From this visual structure, a set of views can now be generated, which allow users to navigate through the display. User interactions can influence the transformation process at different stages. Users can adjust their view on the data, change the visual structure, or even affect the data transformation. The cyclic arrows in the diagram refer to the fact that the processes involved in the distinct steps are of an iterative nature and can occur repeatedly before the next step follows.

**Data type.** Shneiderman (1996) suggested a taxonomy for information visualization designs built on data type and task, the type by task taxonomy (TTT). He distinguished seven data types. High-level abstractions and specific data-types are treated as subordinates of the types presented. In this model, Shneiderman assumes that all data in information space are collections of items, where items have multiple attributes.

- *1-dimensional*
  Text files and alphanumeric list of names
- *2-dimensional*
  Geographic map or book layout
- *3-dimensional*
  Real world objects and chemical molecules
- *Temporal*
  Time-series and scientific measurement rows
- *Multi-dimensional*
  Relational database content
- *Tree*
  Structured data collections with hierarchy constraints
- *Network*
  Structured object sets which do not apply to tree constraints.

Shneiderman (1996) deployed the type taxonomy to sort-out research prototypes of that time and point towards new opportunities. He himself considered the classification incomplete and forecast that upcoming applications would require novel and, respectively, specialized data structures.

A variety of consecutive taxonomies proposed extensions to the TTT, but were never as widely adopted as Shneiderman's work. In his summary, Keim (2002) discards 3-dimensional data and temporal data as data-types on their own. In contrast to the TTT, temporal data is a peculiarity of one-dimensional data. Text in turn becomes promoted a data type, whereas Shneiderman considered it one-dimensional. Hypertext joins text. Central information objects in text as well as in hypertext are documents. As outlined in (Keim, 2002), documents themselves are not atomic but, more often than not, internally are of complex structure and most standard visualization techniques cannot be applied right away. First a transformation of the text data into description vectors is necessary. In the last decade, hypertext has certainly become a widely available and significant data repository. Yet, as hypertext can be considered a directed network structure of documents, considering it a data-type opens up ambiguity in the model. Extending the TTT, Keim introduces software and algorithms as new data types that could be visualized.

**Visual representations.** Visual representations, in general, are structures for expressing knowledge. Long before computer technology emerged, visualizations were well-established and widely used. In their empirical study Lohse et al. (1994) investigate how people classify two-dimensional visual representations into meaningful categories. From this survey, a structural classification of visual representations became apparent.

- *Graphs* encode quantitative information using position and magnitude of geometric objects; graphs typically deploy a Cartesian coordinate or polar coordinate sys-tem
- *Tables* are an arrangement of words, numbers, signs, or combinations of them to exhibit a set of facts or relationships in a compact format
  - *Graphical tables* use color to encode numerical data
  - *Numerical tables* show numeric data in text format
- *Time charts* display temporal data; they differ from tables in their emphasis on temporal data
- *Network charts* show the relationships among components; symbols indicate the presence or absence of components; correspondences among the components are shown by lines, arrows, proximity, similarity, or containment
- *Diagrams*
  - *Structure diagrams* are a static description of a physical object
  - *Process diagrams* describe the interrelationships and processes associated with physical objects
- *Maps* are symbolic representations of physical geography; maps depict geographic locations of particular features using symbols or lettering
  - *Cartograms* are spatial maps that show quantitative data
  - *Chloropleths* use color, grey scale, or texture to code areas of equal value
  - *Isopleths* use lines to join points with the same quantity of value
  - *Dot maps* use points or symbols to show the location of individual points on a map
  - *Flow maps* show direction of movement by the number, width, and direction of lines and arrows
- *Icons* impart a single interpretation or meaning for a picture; each icon provides a unique label for a visual representation
- *Photo-realistic pictures* are realistic images of an object or scene.

The visual artifacts under examination originated from the domain of static, two-dimensional graphic representations. Hence, no statement can be made to what extent the classification also covers three-dimensional or interactive displays. Lohse et al. (1994) themselves mention as a caveat that their classification schema structures the domain of visual representations at a high level. As yet, no deep, hierarchical structures within clusters have been identified. Moreover, the study focuses on perceived similarity among the visual artifacts that were inspected. Instead, a classification must represent structure that is used by people in interpreting graphs.

**Visualization techniques.** The classification identified by Lohse et al. (1994) distinguishes itself by clear terminology. Common terms like diagrams, or specialized terms, like chloropleths, indicate classes of visual representations that deploy a well-defined set of visualization techniques that are already established. More importantly, visual representations often get associated with scenarios in which they are deployed regularly.

Keim (2002) concentrates on the design of the visual environment and suggests a classification of visualization techniques that takes into consideration recent developments in information visualization.

- *Standard 2D/3D displays* deploy traditional visual encodings.
- *Geometrically transformed displays* find appealing and useful geometric transformations of visualizations of multidimensional data sets.
- *Icon-based displays* map the attribute values of a multidimensional data item to the features of an icon.
- *Dense pixel displays* map each dimension value to a colored pixel and group the pixels belonging to each dimension into adjacent areas.
- *Stacked displays* present data partitioned in a hierarchical fashion.

Wiss and Carr (1998) intuitively grouped information visualization designs according to their presentation.

- *Node-link style designs* typically support networks and tree data types.
- *Raised surface designs* display information on surfaces (horizontal or vertical) that gets distorted.
- *Information landscapes* support a variety of data types; they all share 2.5D appearance with information plotted as shapes on a surface.
- *Other designs* that do not fall into previous classes.

The classifications of Keim (2002) as well as of Wiss and Carr (1998) describe high-level procedures for the construction of visual environments.

All classifications of visualization techniques examined are concerned with visual attributes like color (texture, shading), shape, size, position, and (semantic symbols). In dynamic systems, however, time can serve as additional dimension for display. With this approach, all visual attributes can alter in time. The change of position during time is also known as animation.

In the last decades, a large number of novel information visualization techniques have been developed. Good overviews of the approaches can be found in a number of recent books (Card et al., 1999; Ware, 2000; Spence, 2000).

**Tasks.** Bundled with the type taxonomy, Shneiderman (1996) enumerated seven tasks users could perform on the data. Complex tasks, e.g. focus & context, can be described as a combination of tasks presented, in this case overview, relate, and zoom.

- *Overview* Gain an overview of the entire collection.
- *Zoom* Enlarge items of interest.
- *Filter* Filter out uninteresting items.
- *Details on demand* Select an item or group and get details when needed.
- *Relate* View relationships among items.
- *History* Keep a history of actions to support undo, replay, and progressive refinement.
- *Extract* Take out sub-collections of data or history to save and communicate.

**Interaction.** The information visualization process of transforming data into visual representations is a one-way street unless the human perceiver is given the opportunity to intervene. Human interaction completes the loop between visual forms and control of the visualization process. It includes controlling the mappings performed in the visualization process (Card et al., 1999): data transformations, visual mappings, and view transformations.

Although interactive techniques and metaphors differ in design, Chuah and Roth (1996) have identified primitive interactive components visualization systems have in common. Composing these primitives can model the complex behavior of visualization system user-interfaces at the semantic level of design. The functional classification distinguishes between three main types of basic visualization interactions. Each main type ramifies to a hierarchy of more specific interaction types.

- *Graphical operations* affect the graphical representation of data.
  - *Encode data* operations modify visual mappings (Fig. 1).
    - *Change mapping* operations alter existing or create new mappings between data and visual representations.
    - *Transform mapping* operations manipulate the encoding range of mappings, allowing the magnification of differences between values or separate sets of objects.
  - *Set graphical value* operations alter visual representations of selected entities by directly specifying the new value. Hence, the appearance of the affected visual objects no longer solely depends on the underlying data.
    - *Constant* operations set the graphical attribute to a constant.
    - *Graphical transform* operations determine the values of visual attributes through formulas.
  - *Manipulate objects* operations treat graphical objects as building blocks and modify the visual scene independent of underlying data and mappings.
    - *Copy* operations instantiate new graphical objects.
    - *Delete* operations remove graphical objects from the visual representation.
- *Set operations* refer to all those operations that act on or form sets. Sets provide users with the capability to collect and assemble objects that belong together. The underlying data gets enriched with new classification information.
  - *Create set* operations establish new object sets.
    - *Enumerate* operations let users individually pick objects to accommodate from the visualization.
    - *Express membership* operations express conditions for set membership through a formula or constraints. All objects that meet these criteria are automatically added to the set in bulk.
  - *Delete set* operations dissolve sets whereas objects formerly included persist.
  - *Summarize set* operations perform aggregation operations on set members.
- *Data operations* directly affect the data presented by the visualization. In contrast to all other interactions, data operations promote visualizations. Visualizations become a means not only to retrieve but also to input and change data.
  - *Add* operations create new data elements.
  - *Delete* operations destroy data elements.
  - *Derived attributes* operations augment data with new attributes.

Compared to the reference model for visualization (Fig. 1), this classification sprawls beyond the traditional limits of information visualization. In addition to presenting, retrieving, and exploring data, data operations also allow the manipulation of underlying (raw) data. This feature is required for visualizations in order to grow to full-fledged application system user-interfaces.

Set operations also contribute to this development. Although not in the center of the information visualization process, set operations reflect today's code of practice. Theus (2003) adds that setting up complex selection sets usually is achieved by step-wise refinement. Boolean operators combine subsets derived by enumeration or membership rules, creating unions, intersections, and complements.

Interactions to geometrically navigate within the view presented are not considered in that work.

**View transformations.** The visual mapping process results in graphical structures that represent information. In a final step, views render these graphical structures and make them accessible to the human perceiver, on computer screens, for example. View transformations specify graphical parameters that influence the view such as position, scaling, and clipping. Varying view transformations can reveal more information from one and the same graphical structure than static visualizations possibly could. Card, Mackinlay and Shneiderman (1999) distinguish three common view transformations.

- *Location probes* are view transformations that expose additional information based on the position within the graphical structure. When triggered by the human perceiver, location probes could also be referred to as details-on-demand. For display, location probes can either augment the visual structure in the selected region or create additional views.
- *Viewpoint controls* are pure geometrical transformations to zoom, pan, and clip the viewpoint.
- *Distortion techniques* help to maintain orientation during the exploration process (Keim, 2002). Meanwhile the focus is displayed in great detail, the surrounding context remains visible. Distortion techniques graphically transform the visual structure to render focus and context combined within one single view.

Since location probes can, and often do, result in additional views, it is critical to classify them as view transformations. Moreover, location probes also do influence the visual mapping, in case they trigger the enrichment of the visual structure for details-on-demand.

Scales, as introduced by Theus (2003), encompass location probes and viewpoint controls. The former is referred to as logical zoom, whereas the pure graphical operational performed by viewpoint controls is a simple zoom.

Leung and Apperley (1994) introduce transformation and magnification functions for various distortion-oriented presentation techniques. Different classes of functions refine the classification of distortion view transformations.

**Multiple view coordination.** Multiple view systems "use two or more distinct views to support the investigation of a single conceptual entity." (Wang Baldonado, Woodruff & Kuchinsky, 2000) To fully exploit the potential of multiple views, sophisticated coordination mechanisms between views are required.

- *Navigational slaving*
  Movements in one view are automatically propagated to another view.
- *Linking*
  Connects data in one view with data in another view.
- *Brushing*
  Corresponding data items in different views are highlighted simultaneously.

Views are distinct, if they reveal dissimilar aspects of the conceptual entity presented. Roberts (2000) identified three ways in which multiple views may be formed according to stages in the information visualization process comparable to the reference model (Fig. 1). Multiple views from the filter level branch during the data mapping step. Multiple views from different mappings emanate from varying visual mappings. Finally, display-level multiple views arise due to altering the viewport or projection specification.

Multiple views perfectly join with the reference model (Fig. 1), which did consider sequential presentation of views, but no coordination of parallel views.

Theus (2003) reports about the most common use of multiple views.

**Cognition.** By definition, the purpose of information visualization is to "communicate properties of information to a human". The research on information visualization must not stop at producing and designing visualization but must also consider how visualizations affect the human observer. Wiss & Carr (1998) propose a framework for classification of 3D information visualization designs based on three cognitive aspects.

- *Attention* denotes how designs draw attention to certain elements of the visualized data.
  Focus on certain elements of the visualized information. Differences in visual appearance, movement, location, and metaphors can be used to attract human attentions.
- *Abstraction* indicates how designs support information structuring and information hiding.
  Clustering or grouping parts of the information to form higher-level elements.
- *Affordances* measure how designs show to the users what they can do with them.
  The visual cues that a visual element gives to indicate what can be done with it.

Any of these aspects allow for a multitude of peculiarities. A survey revealed that information visualization systems have come up with a variety of solutions in order to guide user attention, abstract from complex data and indicate available functionality and interaction modes. Introducing their solutions as second level classes would turn the current framework of aspects into a classification.

**Information visualization operating steps.** The data state reference model (Chi, 2000) describes visualization techniques with a focus on data and its transformations. The model breaks down the information visualization process into four data stages:

value, analytical abstraction, visualization abstraction, and view. Three types of data transformation operators carry over into states.

- *Data transformation* operations generate some form of analytical abstraction from the raw data values.
- *Visualization transformation* operations further transform analytical abstractions into visualizable content. The visualization abstractions resulting from this are not visual structures yet.
- *Visual mapping transformation* operations map visualizable content into graphical structures.

Another four types of operators cater for data transformation within data stages. Based on the data state model, Chi decomposed the data processing pipelines of visualization techniques and identified operating steps they share.

## 3 Information Visualization Model Consolidation

With our approach, we do not intend to substitute information visualization models and classifications that have evolved so far. Instead, best-of-breed will be selected and combined into one consolidated formal model describing information visualization.

### 3.1 Information Visualization Model Space

All the classification models presented describe selected subsets of the complex area of information visualization. Our attempt to arrive at a consolidated model for information visualization starts out with the analysis of what areas these discrete models cover and how they are mutually related (Fig. 2). To answer that question, we locate information-visualization models within *model space* for information visualization. There are two axes that span model space. The first dimension reflects the processing pipeline for (information) visualizations as introduced by the reference model for visualization (Fig. 1). Roughly speaking, three sections subdivide this pipeline. Beginning with the data section, data is transformed and mapped into graphical objects in the visualization section. Of course, models describing data properties, for example, are located to the left whereas multiple views and their coordination cover the area from the middle to the right. The second dimension expresses dependencies between models as well as the level of abstraction from the actual task of handling (computer) data. On the lowest level, models deal with data properties and visual attributes, whereas at the upper levels, models such as cognition abstract away from implementation details. Upper level models depend on their subordinates. The absence of visual objects and their properties would render talking about cognition futile.

Of course, as information visualization model space lacks metrics, positions and borderlines get blurred. So far, the diagram reflects our subjective assessment. Furthermore, drawing rectangles is a simplification. More often than not, single models

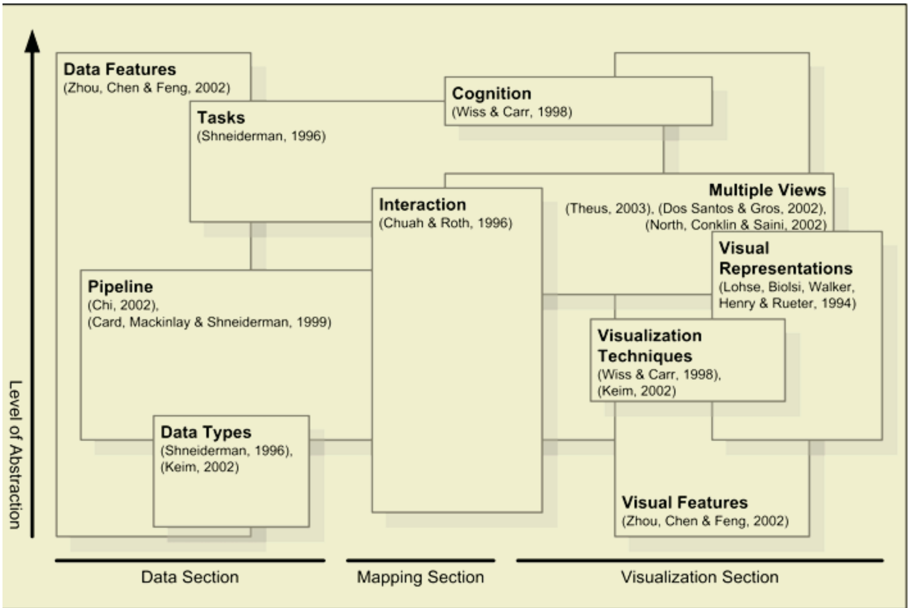


**Fig. 2.** Interrelationship of information visualization models in information visualization model space

do not handle all aspects at one constant level of abstraction and vice versa. This holds true especially for substantial models. Hence, the areas in the diagram depict an approximation of the real state of affairs.

### 3.2 Coverage and Ambiguity

The first overview reveals that there is little white space in the diagram. Judging from that, the extant models in total cover nearly all facets of information visualization as we know it today.

The frayed right side of the visualization section indicates that information visualization model space has no clearly marked border in this direction. Multiple views, visual representations, cognition, and interaction not only apply to information visualization exclusively. Partially, these models belong to visualization in general. From our point of view, visualization model space begins in the visualization section and extends beyond the diagram.

The next observation is that rectangles in the diagram overlap. If this occurs within one section, the models involved compete. Such conflict can be observed, for example, between data types, as introduced by Shneiderman's (1996) TTT, and the data features invented by Zhou et al. (2002). Sorting out the differences and matching concepts are the anticipated tedious tasks required in order to arrive at a joint model. The above presentation of information visualization models discusses corresponding models. Note that the collection of models portrays selected samples. Less important items have already been omitted.

Sections cannot always be clearly separated without ambiguity. Cross-section overlapping arises when one and the same phenomenon of information visualization is covered by various models starting out from different perspectives. For instance, interaction and the processing pipeline are closely interwoven. From the standpoint of the reference model, view transformations are modifications that are likely to be triggered by human interaction. Conversely, interaction claims that location probes and viewpoint controls are their terrain, and terms them interactive filtering, interactive zooming, interactive distortion, and interactive linking and brushing.

### 3.3 Quality and Level of Granularity

As the diagram suggests, the area of information visualization has been thoroughly researched and only few white spaces remain. Yet the stake the various models claim reflects neither the model quality nor its level of detail. There are always two sides to quality: correctness and completeness. Before they can be integrated into the coherent model, extant models need to be assessed with care. More easy to judge is the model's level of granularity. Classification systems vary in how detailed a way they have been conceived. Generally, coarse models leave space for alternatives and variations, whereas in depth models provide better guidance. To illustrate the difference, the interaction model with three hierarchy levels of classes is far more detailed than the data types according to the TTT. Then again, not all facets of information visualization share the same level of complexity. It is natural that different areas feature different numbers of classes.

## 4 Information Visualization Modeling Language

Current practice in information technology favors the use of formal languages as representation formalisms which abstract away from details of specific realization. The information visualization modeling language enables the declarative description of an information visualization need or solution in preference to describing the steps required in order to realize the visualization process. It is a formal language; it has a set of strings which can be derived from a (formal) grammar consisting of a deductive system of axioms and inference rules (Partee, ter Meulen & Wall, 1990). We give the term information visualization modeling language *blueprint* to the formal description of an information visualization technique or application expressed by the language. A blueprint is composed of a number of sections. Blueprint sections are legal combinations of language elements derived from the grammar.

Conceiving the information visualization modeling language may follow two simple rules of thumb. First, concepts identified within the model constitute the vocabulary. Secondly, relationships between concepts determine the grammar. Presumably, however, relationships from the model will also contribute to the language vocabulary. The information visualization modeling language will constitute a specific encoding of the consolidated information visualization model. In order to be useful, its design has to meet requirements for both computation and application.

### 4.1 Computational Desiderata

The information visualization modeling language (IVML) carries knowledge about information visualization within its schema. Moreover, information visualizations denoted in the language are formal structures which represent knowledge about information visualization techniques, applications, and requirements, respectively. Hence, the information visualization modeling language can be considered a meaning representation language. Meaning representation languages need to meet a number of practical computational requirements (Jurafsky & Martin, 2000).

*Verifiability* is the most basic requirement for a meaning representation: “it must be possible to use the representation to determine the relationship between the meaning of a sentence and the world as we know it.” In the case of the IVML, it can (say) describe information visualization techniques and data types these techniques are capable of displaying. These descriptions establish knowledge. Demands for visualization of data of a specific type can be considered a question expressed in IVML. If there is no visualization technique that can handle the requested data type, matching will fail. In general, sentences can have different meanings depending on the circumstances in which they are uttered. Since the IVML is intended to be the means we reason about and act upon, it is critical that blueprint sections expressed in the language (analogous to natural language sentences) have single unambiguous interpretations. The IVML is required to be an *unambiguous representation*. Conversely, distinct sentences in general may have the same meaning. Such a situation is highly problematic, since it hinders verification and adds complexity to reasoning. Therefore, the IVML should follow the doctrine of *canonical form*: Sentences that mean the same thing should have the same representation. More complex requests cannot be answered solely on the basis of verification and canonical form. Let’s agree that whilst traditional diagrams in general are suitable for presentation purposes, they are not a good choice to pursue data exploration. Pie charts belong to this class of traditional visualization techniques. To meet the demand for visualization of data for presentation purposes using pie charts, *inference* is required. It must be possible to draw conclusions about propositions that are not explicitly represented, but are nevertheless logically derivable from the knowledge available. Finally, in order to be useful, the IVML must be *expressive* enough to treat a wide range of the subject matter of information visualization. But, since research in this area is ongoing, the IVML cannot be expected to be complete.

### 4.2 Applicational Desiderata

By analogy with design criteria that underlie related modeling languages (Web3D Consortium, 1997), the information visualization modeling language should meet a set of requirements in order to be useful in application.

Information visualization is a multifaceted subject matter. The formal description of information visualization techniques and applications using the IVML will be accordingly complex. *Composability* provides the ability to use and combine information visualization objects, like data sources, mapping formulas, or view definitions, within an IVML application and thus allows reusability. Depending on the application, the complete set of constructs is not always required. In a single-view application, for example, multiple-view coordination is pointless. The design of the IVML must permit the omission of constructs which are not essential for the given situation. The notion of language constructs which are independent by design is known as *orthogonality*. Since the IVML is anticipated not to cover all future inventions in the area of information visualization, the language has to be *extensible*, allowing the introduction of new concepts. Wherever concepts are missing in the language, *bypass*es help to fill the gaps with alternative solutions. Bypasses also stand in when IVML design does not meet particular requirements. In the case of parsers interpreting the IVML in order to render information visualizations, the bypass addresses purpose-built implementations. The IVML needs to be *authorable*: Computer programs must be capable of creating, editing, and maintaining IVML files, as well as automatic translation programs for converting related data into IVML. More generally, the language must be *capable of implementation* on a wide range of systems. Considering the implementation of software systems, language design must foster the development of scalable high-*performance* implementations. Finally, IVML must *scale* and enable arbitrarily large dynamic information visualization applications.

## 5 Modeled Information Visualizations in Education

The characteristics of the IVML can greatly contribute to the field of education. Today, mapping tools for visualizing conceptual knowledge are broadly available and so are mapping tools for visualizing content knowledge. Out of these mapping tools, a large number allow for designing and utilizing visualizations in a quick and efficient manner. Modern, computer-based mapping tools are capable of handling both content and conceptual knowledge at the same time. In the majority of cases, however, mapping tools are not capable of externalizing the knowledge implicitly embedded in the design of the visualization of content knowledge. This was, however, essential for the application of mapping tools in education. Imagine tutors designing lectures or self-driven students generating individual maps while learning. The tutor notes the created map so that students later can recall and use it. Students in turn conserve their individual map. Or, for sharing their experience, communicate it to peers. In all these scenarios, the modeling language captures structure, appearance, behavior, and functionality of the visualization. In fact, IVML stores knowledge about the mappings.

### 5.1 Peer Students Scenario

Imagine a student investigating in a self-regulated fashion, being engaged in an information retrieval dialogue with a computer-based interactive information visualization system, seeking to meet an information need he cannot fully specify. Hence, it is impossible for him to formulate a question and have the system answer in a targeted way. Instead, the dialogue is of an exploratory nature. During a series of iterative steps the student learns about the data source, locates relevant information, and refines his information need. This process is put into practice by human actions demanding the system to adapt in return. Beginning with an initial setup, interactions manipulate data transformations, visual mappings, and view transformations. Finally, if the dialogue succeeds, the student will have come to a relevant data set answering his information needs.

So far, the mapping tool and its functionality supported the student in primarily one out of a variety of knowledge management processes (Tergan, 2003): The student *localize*d content knowledge and the corresponding knowledge resources. But there is more to it. At the end of the dialogue, the student will not only have come to a relevant data set answering his information needs, but moreover end up with an information visualization application tailored to the task performed. During the dialogue, knowledge has been *generate*d. On the one hand, the student gained new insights. On the other hand, the mapping tool configuration reflects how the knowledge has been derived and justifies the new knowledge.

Imagine the system was able to export its final state as a blueprint. The information visualization modeling language would then be deployed to formally *represent* the knowledge about the information visualization technique that has evolved, allowing it to be noted down (electronically). Usually, only content retrieved is retained as a result of the dialogue, discarding the history and the supporting tool's setup. With the various blueprint sections, all these facets of the information retrieval dialogue can be preserved. Furthermore, the knowledge captured using the information visualization modeling language can be re*use*d in similar tasks or applied to diverse data sources. With the blueprint the information visualization technique can be *communicate*d in its entirety to third parties, particularly to peer students.

## 5.2 Tutor Scenario

Imagine the lecturer creating a new tutorial. Unlike in traditional lectures, students are expected to study in a self-regulated fashion, investigating given resources utilizing an interactive information visualization system. Basically, the tutor needs to decide what resources to base the tutorial on. More challengingly, he must come to a decision what presentation styles and what investigation tools are most appropriate to support the students in achieving the intended learning outcome. In the case of information visualization techniques, this design task implies the definition of data transformations, visual mappings, and view transformations. Depending on the learning outcome, the tutor will choose proper mappings, multiple view constellations and available interactions in order to enable the students to reveal patterns, clusters, gaps, or outliers, for instance. The resulting information visualization techniques then get distributed to the students as teaching aids who in turn apply them to the resource data and, hopefully, gain new insights.

Again, the information visualization blueprint gets deployed to store and distribute all facets of information visualization applications, thus *represent*ing and *communicat*ing knowledge about the information visualization technique and its application. Beyond applying blueprints prepared by the tutor, students may compare the encoded experts' knowledge with blueprints they created themselves and *evaluate* their own knowledge this way. The information visualization modeling language also may support the tutor in authoring appropriate blueprints. Blueprint sections may capture the tasks and goals established visualization techniques serve best. With this knowledge, the language may help the tutor to select and adapt appropriate techniques and foster the re*use* of knowledge this way.

## 6 Summary and Conclusion

This article outlines our approach towards the information visualization modeling language (IVML). To lay a sound foundation, we survey the state-of-the-art of information visualization, assess the coverage and relationships between extant models, and identify potential obstacles in the process of setting up an integrated formal model that reflects the concepts and relationships of information visualization as it is understood today. Finally, we present computational requirements as well as those imposed by the application the information visualization modeling language has to fulfill.

The survey focuses on work that devised classification schemas. To assess which facets of information visualization these discrete models cover and how they are mutually related, we established the notion of information visualization modeling space. The analysis suggests three findings. First, the extant models in total cover nearly all facets of information visualization as we know it today. Secondly, areas of information visualization model space are described by rival models, leading to ambiguity. Third, the models vary in the level of detail in which they have been worked out.

In order to mutually open-up the treasure chests of visualization techniques for knowledge and information, information visualization and knowledge visualization must share, at least to a significant amount, a joint visualization model. Achievements in information visualization then could get applied to knowledge and vice versa.

The information visualization modeling language constitutes a specific encoding of the consolidated information visualization model. Its design has to meet requirements for both computation and application.

Two scenarios suggest how the information visualization modeling language could contribute to the field of education, supporting students studying in a self-regulated fashion. The benefits arise from the language's capability to formally represent, note, preserve, and communicate structure, appearance, behavior, and functionality of information visualization techniques and their applications. In combination with interactive mapping tools, the modeling language assists students and tutors in the evaluation, localization, generation, representation, communication, as well as the use of knowledge.

## References

Card, S., Mackinlay, J., & Shneiderman, B. (1999). Information Visualization. In: S. Card, J. Mackinlay, & B. Shneiderman (Eds.): *Readings in Information Visualization* (pp. 1-34). San Francisco, CA: Morgan Kaufmann.

Chi, E.H. (2000). A Taxonomy of Visualization Techniques Using the Data State Reference Model. In S.F. Roth (Ed.), *Proceedings of the IEEE Symposium on Information Visualization (InfoVis 2000)* (pp. 69-76). Washington, DC: IEEE Computer Society Press.

Chuah, M.C., & Roth, S.F. (1996). On the Semantics of Interactive Visualizations. In S. Card, S. Eick, & N. Gersham (Eds.), *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'96)* (pp. 29-36). Washington, DC: IEEE Computer Society Press.

Jurafsky, D., & Martin, J. (2000). *Speech and Language Processing*. Upper Saddle River: Prentice Hall.

Keim, D.A. (2002). Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, *8*(1), 1-8.

Leung, Y., & Apperley, M. (1994). A Review and Taxonomy of Distortion-Oriented Presentation Techniques. *ACM Transactions on Computer-Human Interaction*, *1*(2), 126ff.

Lohse, G.L., Biolsi, K., Walker, N., & Rueter, H.H. (1994). A Classification of Visual Representations. *Communications of the ACM*, *37*(12), 36-49.

Partee, B., Ter Meulen, A., & Wall, R. (1990). *Mathematical Methods in Linguistics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Roberts, J.C. (2000). Multiple-View and Multiform Visualization. In R.F. Erbacher, P.C. Chen, J.C. Roberts, & C.M. Wittenbrink (Eds.), *Visual Data Exploration and Analysis VII, Proceedings of SPIE*, Volume 3960 (pp. 176-185). Bellingham, WA: SPIE The International Society for Optical Engineering.

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the IEEE Symposium on Visual Languages 1996* (VL '96) (pp. 336-343). Washington, DC: IEEE Computer Society Press.

Spence, R. (2000). Information Visualization. Harlow, England: Addison-Wesley.

Tergan, S.-O. (2003). Managing Knowledge with Computer-Based Mapping Tools. In D. Lassner, & C. McNaught (Eds.), *Proceedings of the World Conference on Educational Multimedia, Hypermedia & Telecommunication (ED-Media 2003)* (pp. 2514-2517). Charlottesville, Country: AACE Press.

Theus, M. (2003). Navigating Data – Selections, Scales, Multiples. In C. Stephanidis, & J. Jacko (Eds.), *Human-Computer Interaction - Theory and Practice (Part II). Proceedings of HCI International 2003 (HCII 2003)* (pp. 1323-1327). Mahwah, NJ: Lawrence Erlbaum Associates.

Wang Baldonado, M.Q., Woodruff, A., & Kuchinsky, A. (2000). Guidelines for Using Multiple Views in Information Visualization. In V. Di Gesù, S. Levialdi, & L. Tarantino (Eds.), *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI'00)* (pp. 110-119). New York: ACM Press.

Ware, C. (2000). Information Visualization: Perception for Design. San Francisco, CA: Morgan Kaufmann.

Web3D Consortium (1997). Information Technology - Computer Graphics and Image Processing - The Virtual Reality Modeling Language (VRML) - Part 1: Functional Specification and UTF-8 Encoding (VRML97), ISO/IEC 14772-1:1997.

Wiss, U., & Carr, D. (1998). *A Cognitive Classification Framework for 3-Dimensional Information Visualization*. Research report LTU-TR-1998/4-SE. Luleå, Sweden: Luleå University of Technology.

Zhou, M.X., Chen, M., & Feng, Y. (2002). Building a Visual Database for Example-based Graphics Generation. In P. Ch. Wong, & K. Andrews (Eds.), *Proceedings of the IEEE Symposium on Information Visualization 2002 (InfoVis'02)* (pp. 23-30). Boston, MA: IEEE Computer Society Press.