Multimedia Tools and Applications (2021) 80:12541–12579
https://doi.org/10.1007/s11042-020-10306-9

A proposed UML-based common model for information visualization systems

Cristian Moral[1] · Angélica de Antonio[1] · Xavier Ferre[1] · Jaime Ramirez[1]

Received: 20 May 2020 / Revised: 28 September 2020 / Accepted: 22 December 2020 /
Published online: 12 January 2021

Abstract
Nowadays, visualization of large amounts of data has become a key issue for processing data in many fields. This paper presents a conceptual model for describing information visualization systems. This conceptual model has been represented using the Unified Modeling Language, therefore it takes advantage of the understandability, unambiguity, flexibility, and adaptability provided by this notation. The proposed model outlines concepts such as visualization workspace, device, visual representation and its features, information visualization task and its effects on visualization, inter-item relationship, etc. The main applications of the proposed model are: (1) to guide the development of a new visualization system by specifying which are the main concepts that have to be considered and how they relate to each other; (2) to characterize or describe an already existing visualization system; and (3) to thoroughly compare existing visualization systems. To illustrate the use and applications of the model, several examples of information visualization systems are provided.

Keywords Data visualization · Human computer interaction · Information visualization modelling · User interfaces

✉ Cristian Moral
cmoral@fi.upm.es

Angélica de Antonio
angelica@fi.upm.es

Xavier Ferre
xavier.ferre@upm.es

Jaime Ramirez
jramirez@fi.upm.es

[1] Universidad Politécnica de Madrid, E.T.S de Ingenieros Informáticos, Madrid, Spain

1 Introduction

In the last decades, processing large amounts of data has become very common in multiple fields, like economics, health care, biology, sociology, information retrieval, etc. In fact, big data and data mining solutions have undergone a great growth in the last years, motivated by the great increase in data generated by countless devices and digital services, for example, smartphones and social network applications, but also thanks to the rapid increase in processing capabilities of computers. It is essential not only to process this huge amount of data, but also to visualize them in a proper manner so that they become understandable for and processable by the user. As Shneiderman stated: "*Abstract information visualization has the power to reveal patterns, clusters, gaps, or outliers in statistical data, stock-market trades, computer directories, or document collections*" [74].

The main characteristic of information visualization, regardless of its purpose, is to "*use [...] computer-supported, interactive, visual representations of abstract data to amplify cognition*" [19]. Moral et al. [52] conducted a qualitative study of the information-seeking processes performed by computer science researchers and then concluded that information visualization is essential to help researchers analysing hundreds or thousands of documents in order to find those that fulfil their information needs [51]. This is due to the inner perception and cognitive capacities of humans to recognize, process, and analyse visual representation that allow inferring complex information and resolving logical problems [10, 86].

In this paper, we propose to properly conceptualize how information has to be visualized to facilitate information-seeking. To reach this goal, we have identified which are the main aspects that have to be considered when designing an information visualization system. Then, we have represented them through a UML-based conceptual model. More precisely, we have used part of the notation for class diagrams proposed in the Unified Modeling Language (UML) [58] to construct our pictorial conceptual model.

UML is a well-known and standardized language that allows creating a model with a high degree of understandability and unambiguity as the notation, its use and its meaning are well defined. UML also provides a high degree of flexibility and adaptability as it allows representing the concept of inheritance, where properties of a given concept are inherited by the sub-concepts that are a subtype of the first concept. This standard notation is widely used mainly in software engineering [81], but also in other fields like design of simulation models [66].

In the latter case, conceptual models are used to guide the design process, considering the elements that are relevant from user's perspective. In this direction, Sharp, Preece & Rogers state that *"most interface applications are actually based on well-established conceptual models"* [63]. Besides, a pictorial model, like a UML-based model, provides the information in a visual form that can serve to more clearly transmit the relevant information to stakeholders due to the excellent human capacity to process visual information [35]. As it will be explained later in the paper, this can be especially useful to easily and thoroughly describe and/or compare existing information visualization systems, and to guide the design of new ones.

Even though the model has been developed to be general, three examples of real information visualization systems created to visualize a specific type of data (papers, books, theses, etc.) in a specific context (research-oriented information-seeking), are utilized along the paper in order to facilitate the understanding of the model:

- Google Scholar [30] (see Fig. 4) is a well-known 1D meta-search engine that returns results obtained from different sources, like digital libraries from research editorials, or patents databases. In addition to a list of publications, it also provides several information related to them, like who are their authors, where and when they have been published, who has cited them, or a suggestion of related publications.
- Calimaco [24–26] (see Fig. 5 and Fig. 6) is an information visualization system developed in a 3D virtual environment. Documents, which are represented as 3D spheres, are manipulated by the user through mid-air gestures. The system allows to automatically cluster them according to their content and encodes their thematic similarity through the spatial position of the spheres: the closer the spheres are, the more similar the documents they represent are. Besides, spheres representing documents that belong to the same cluster have the same colour.
- 3D Explorer [60, 61] (see Fig. 2 and Fig. 3) makes use of a multi-axis system to spatially organize the documents' representations according to the topic(s) they address. Each axis represents a topic and uses the colour to encode the thematic similarity between documents. 3D Explorer, additionally, allows displaying the documents as a coloured alpha-numerical list.

In the remaining of the paper, we first state the motivation for this work and review the literature. Then, we present the methodology used to derive the model. In Section 5, we present the conceptual model of information visualization, and exhaustively explain all the concepts and relationships it contains. Section 6 describes some of the uses that can be given to the model, presenting specific examples. After this, we compare our proposal with other models and frameworks existing in the literature. Finally, we summarize the main contribution of the paper and propose the future work that could be done based on this contribution.

2 Motivation

According to Lu et al. [40], education and training is an unsolved problem in the information visualization field. One of the main deficiencies that has been identified is the need to provide theoretical foundations that help to design and create new visualizations, considering both the user (needs, preferences, abilities, or limitations) and the technology (through guides, principles, models, frameworks, or idioms).

In fact, as pointed out by Jaeschke et al. [34], several authors have proposed models dealing with specific parts of the information visualization, such as the tasks that can be performed [4, 5, 14, 72, 74, 90], the visualization techniques that are available to display information [62, 69, 88], or the visual features used to encode the characteristics of the data being displayed [93], among others. However, even if all these models are interrelated, nobody has proposed a holistic model encompassing all of them.

This gap has been widening in recent years with the explosion of big data, which makes it essential to design efficient visualization systems that allow extracting information from a very large set of data by identifying, for example, patterns [46].

In recent years, a lot of research has been dedicated to model the information visualization process. This effort has mainly focused in describing the pipeline of steps or stages of the process [47, 85]. However, to the best of our knowledge, nobody has provided a conceptual model illustrated through a visual, standard-based, and non-ambiguous representation, which represents all the main concepts and relationships that are relevant in an information visualization activity, but in terms of semantics, not in terms of temporal sequence. Then, even if the tasks allowing the user to interact with the visualization should be described in the model, the temporal dependencies among them are out of the scope of this model.

Our purpose is to complement the existing information visualization process models by proposing a comprehensive conceptual model of the information visualization activity that will help designing, creating, assessing, and comparing information visualization systems. Additionally, we propose using UML to represent the model to provide it with a great flexibility, as it allows to add, modify and delete concepts and relationships to adapt it to new realities, or to include new aspects that may not have been considered. This flexibility is essential to ensure the generalizability, usefulness, and durability of the model in a changing context in continuous evolution.

3 Related works

In the literature, there are dozens of publications explaining how specific information visualization systems, frameworks or toolkits work, like Starlight [67], Polaris [82], prefuse [32], VESTA [50], 3D Medical Visualization System [2], VizAssist [13], Reactive Vega [70, 71], DeepEye [41], VisComposer [48], or VizDSL [54].

However, in this paper we do not intend to present a specific technology or solution, but to propose a general model of the concepts (and their relationships) intervening in an information visualization activity. For that reason, we have reviewed in the literature the works related both to the visual features that have to be considered when designing an information visualization system, and which theoretical information visualization models or frameworks have been already proposed.

3.1 Visual features

During the last decades, information visualization has been extensively studied and it has notably evolved [39]. One of the first authors that laid the foundations of modern information visualization was the French cartographer Jacques Bertin, who, based on his experience, identified some of the essential concepts intervening in an information visualization activity [10]: marks, that is, the graphical objects used to represent data in the visualization, and their positional, temporal, and retinal properties.

The positional property relates to the dimensionality used to illustrate the mark; temporal properties deal with the animation aspects, like mark's movement; and retinal properties encompass all the visual features used to visually encode some information, for example, colour, shape, size, saturation, texture and orientation.

A few years later, Mackinlay [43] added the enclosure and connection properties as a means of illustrating relationships between objects. MacEachren [42] and Healey, Booth & Enns [31] also expanded the types of visual features that can be used to transmit information, including new elements like colour value, hue and saturation, transparency, width, curvature, and flicker, among others.

The way in which marks and their visual features are used to transmit information is very variable. In fact, many authors have stated that visualizing information is a domain-specific activity as there are many aspects that must be considered when designing an information visualization system [4, 62]. These aspects are the type and structure of the data to be displayed [65, 73, 77, 83, 89], the user, in terms of perceptual and cognitive capabilities, but also of preferences and needs [4, 5, 20, 38, 57], and the context [1, 16, 22, 37].

In contrast, in this research we aim at providing a general domain-independent model including all the visual features that may be used in an information visualization system, so that designers may select the subset of them that best meets their needs.

3.2 Process models and frameworks

Regardless of the domain in which an information visualization is performed, there is some consensus about which is the sequence of steps to be followed to properly visualize information. The reference model for visualization of Card, Mackinlay & Shneiderman [19] is one the most recognized, used and adapted in the field [21, 45]. In all cases, process models define a set of stages that are required to transform raw data into visualizable information that can be visually processed by humans. These stages require some operations, first to uniformize and analyse the unorganized initial information, and then to map it into visualization structures with which the users can interact.

Aaltonen & Lehikoinen [1] also adapted this reference model to make it capable of taking into account the context in which the visualization activity is carried out. To do this, they propose to incorporate a set of simple computational-like rules in the data transformation stage so that, depending on the context, raw data is processed one way or another. Additionally, they propose to create several layers of visual structures, so that the visualization varies according to the context.

Irshad et al. [33] have also adapted the data state model of Chi & Riedl to contemplate the design of information visualization systems in Augmented Reality environments. In this case, the "View" and "Control" concepts of the reference model have been expanded to include aspects like the presentation of the information, the physical interaction, and the sharing of the experience in AR environments.

Similarly, Fernandez & Fetais [29] define the visualization process through a procedural pipeline of 5 steps: *"Map to a shape"*, *"Determine the layout"*, *"Determine attribute visualization"*, *"Add interactive options"*, and "*Render the model"*. The paper pays special attention to the interaction with the system, and describes the 10 visualization features that should be manipulable in an information visualization system to allow performing the four main tasks a user should be able to perform: *"Extract more details from an object", "Show less information"*, *"Modify the values of parameters"*, and *"Customize the scene"*. All the design steps of this pipeline (steps one to four) are covered by our model, as it allows describing all the visual features of the elements (including shape and the rest of attributes), the relationships between elements in the layout, and the techniques that allow to interact with the visualization to perform some user tasks.

Munzner [55] has also provided a procedural information visualization model, in this case formed by four nested layers: *"domain problem characterization"*, *"data/operation abstraction design"*, *"encoding/interaction technique design"*, and *"algorithm design"*. The model tries to cover the full process of creating a visualization system for a given domain and focuses on the importance of defining appropriate evaluation mechanisms to assess the validity of the outputs of each level.

This model was later expanded by Meyer et al. [49] who added the concepts of *"blocks"* and *"guidelines"*. A block represents the internal results of a layer, whereas a guideline indicates how the different blocks interrelate, both within- and between-levels.

In contrast to these process models, which focus on describing the sequence of steps that have to be performed to achieve the desired result, the purpose of this research is to determine which are the relevant concepts that intervene in the process, and how they relate without considering temporal issues, such as precedence. In our opinion, both conceptual and process models complement each other, as they provide a different point of view of information visualization.

3.3 Non-process models and frameworks

Some authors have developed non-process information visualization models. Bugajska [15, 16] provides a complete definition of many of the concepts (and their relationships) intervening in information visualization, like the user, the tasks, the visual representations and their visual characteristics, the navigation within the visualization workspace, or the environments in which the visualization is performed, among others. However, this definition is mainly provided in textual form, as she provides a small graphical framework, but it is a high-level model that only includes a few of these concepts and relationships.

Jaeschke et al. [34] tried to combine several existing models in order to create a consolidated way to describe an information visualization system. As a result, they proposed using a standard notation called *"Information Visualization Modelling Language (IVML)"* to describe the concepts implied in information visualization, and their relationships. Nonetheless, their result does not include the language itself, but the requirements that should fulfil the language, like verifiability, unambiguity, expressiveness, extensibility, or scalability. In our case, on the contrary, we aim at providing designers with a finished tool to help them describing, creating, or comparing information visualization systems.

Munzner [56] proposes some guidelines to help designers taking good design decisions when creating a new information visualization system. Her framework is divided into three parts that, according to Munzner, have to be considered when designing a new information visualization system: *"What?"*, which refers to the visual elements that are used to illustrate the data (for example, tables networks, fields and geometry in the case of dataset, or items attributes, links position and grids in the case of individual data); *"Why?"*, which deals with the analysis of the tasks that can be performed in an information visualization system (for example, discover, annotate, lookup or compare) and which are the targets of these tasks (for example, finding a trend in a dataset or similarities between different data); and *"How?"*, which explores which kind of mechanisms uses the information visualization system to allow the user to analyse the data (for example, ordering, selecting, navigating, superposing or filtering a set of data, or mapping some categorical and ordered attributes from the data by using colours, size or shape).

Our model mainly focuses on the two later questions, as it allows defining which are the tasks that can be performed in a system and how the system allows elaborating and analysing the data through a set of visual cues.

Wilkinson [87] does not propose a model as such, but rather provides a set of *"grammatical rules for creating perceivable graphs"* of any kind, and explains *"the meanings of the representative symbols and arrangements we use to display information"*. With these two aspects, Wilkinson covers both the syntax and the semantics of the graphics used in an information visualization system. The purpose of this work is to provide a grammar to guide the design process by providing the theoretical and mathematical base behind some of the aspects that have to be considered when designing an information visualization system, like the scaling, transforming or operating raw data. Wilkinson also explains some aspects related to the visualization itself like how to use the geometry to represent information, or the relevance of the aesthetics when visualizing information.

In contrast to Wilkinson's contribution, our model is a more high-level and visual design tool that focuses on the concepts and relationships themselves, without paying attention to the theory or mathematics behind their representation or manipulation. This allows simplifying the conceptual representation of an information system, and then facilitates its initial conception, its description, and its comparison with other visualization systems. However, we consider that both contributions complement each other, since they both address the design and description of an information visualization system but at different levels of abstraction and with different levels of detail.

Oliveira et al. [59] proposes a meta-model that helps designers selecting the best visualization method based on two aspects: the type of information to be visualized, and the level of detail that is required in the visualization. The meta-model, presented using a Treemap, also makes use of two visualization features (colour and size) to illustrate three aspects: if the visualization is a process or a structure; if the purpose of the visualization is reducing or adding complexity; and if the user wants to have an overview of the information, or wants to dig into the information of individual elements. If the user can specify these three aspects, the meta-model provides a very concrete guideline. However, it does not allow to define the information visualization system itself, nor its specificities, which our model does.

4 Research method

The starting point for this work is the result of a qualitative study performed by Moral et al. [52] to analyse the information-seeking process performed in the computer science research context. The result of this qualitative study was a coding system including 169 codes that were used to code the answers of the 17 researchers that were interviewed or participated in a focus group.

Later, Moral et al. [53] made use of these codes to create a UML-based model to describe all the concepts intervening in the information-seeking process performed by computer science researchers. However, the authors specifically excluded all the concepts related to document storage and management, visualization, and interaction, due to their complexity and the lack of detailed results about these concepts.

Nonetheless, the focus group and interviews performed by Moral et al. [52] made clear the importance of visualization in the information search process, and the lack of a comprehensive conceptual model that can formalize and guide the design of information visualization systems. To fill this gap, this work tries to conceptualize the information visualization activity through a UML-based conceptual model, as it was done for the information-seeking process [53]. The final purpose is to make them complementary.

To achieve this goal, the codes appearing in the coding system appearing in [52] that were directly or indirectly related to the visualization, were taken as the starting point. These codes were all the ones in the "Visualization" category, but also some of those included in other categories like "Workspace", "Exploration", "User tasks", or "Metainformation". All codes categorized as “Visualization” were directly related to the information visualization process, whereas most of the other codes reflected the changes occurring in the visualization after the performance of user tasks.

These codes were used to create an initial model. This model was then incrementally and iteratively improved and extended following two strategies. First, a literature review was carried out (summarized in Section 3) to collect, describe, and organize the concepts related to the visualization of information that appear in relevant works. On the other hand, several information visualization systems were evaluated to characterize their components, functionalities, and features. In both cases, each concept or aspect was analysed and included in the model if considered appropriate. The appropriateness of a concept or aspect was determined in a syndicated way in collaborative work sessions in which researchers with years of experience in the field of research participated. The main criteria that were followed in these sessions were:

- The concept/aspect is directly related to visualization.
- The inclusion of the concept/aspect maintains the coherence and consistency of the whole model.
- The inclusion of the concept/aspect maintains the flexibility and adaptability of the model to make it usable in any context.

Finally, refactoring was performed after including new concepts/aspects to the model, in order to determine if new generalized classes had to be created to group related concepts/aspects.

Table 1 illustrates the UML components that have been used to create the model.

5 Information visualization model

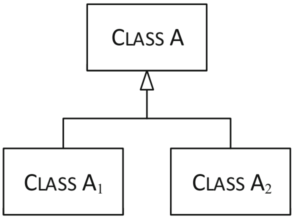
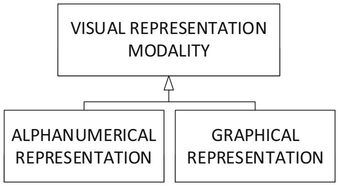
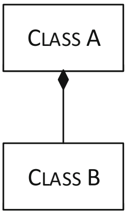
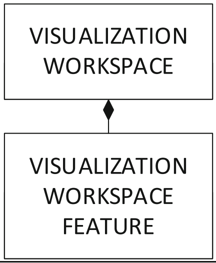
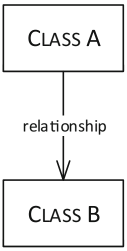
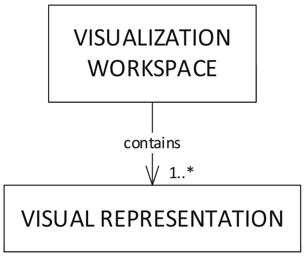
The model presented in this paper, displayed in Fig. 1 and explained in the following sub-sections, has been developed to be general and domain-independent. However, to facilitate the understanding of the model, we also provide examples of a specific domain: document-based information-seeking performed by researchers of any kind. As a result, the general concepts *Item* and *Item Collection* appearing in the model are instantiated in these examples by the specific concepts *Document* and *Collection of documents*, respectively.

To reduce the complexity of the model, some concepts have been duplicated to avoid having lines that cross the entire diagram. These concepts have been coloured to show where they are duplicated (the same background colour means it is the same concept).

5.1 Information visualization activity

An *Information Visualization Activity* can be defined as a specific use of an information visualization system. This use is unique as it takes place in a given context, and is performed by a given user, with a certain mood, characteristics and with a given information need, who chooses some of the visualization workspaces available in the system to fulfil his/her information-seeking purpose. This election establishes which are the input and output devices that have to be manipulated, which type of visual

Table 1 UML elements used in the model

Name	Representation	Description	Example
Class		Class A represents a concept.	*View* is a concept. 
Generalization		Classes A_1 and A_2 are two types of class A.	*Alphanumerical Representation* and *Graphical Representation* are two types of *Visual Representation Modality*. 
Composition		An instance of class A “has” one or more instances of class B.	A *Visualization Workspace* has at least one *Visualization Workspace Feature*. 
Association		Class A is related to Class B as defined by “relationship” .	A *Visualization Workspace contains one or more Visual Representations* 

representations are going to be used to display the information, and which information visualization tasks the user is going to be able to perform to meet his/her target.

A device can simultaneously display more than one visualization workspace, for example by splitting its screen into different areas, one per visualization workspace. Similarly, the

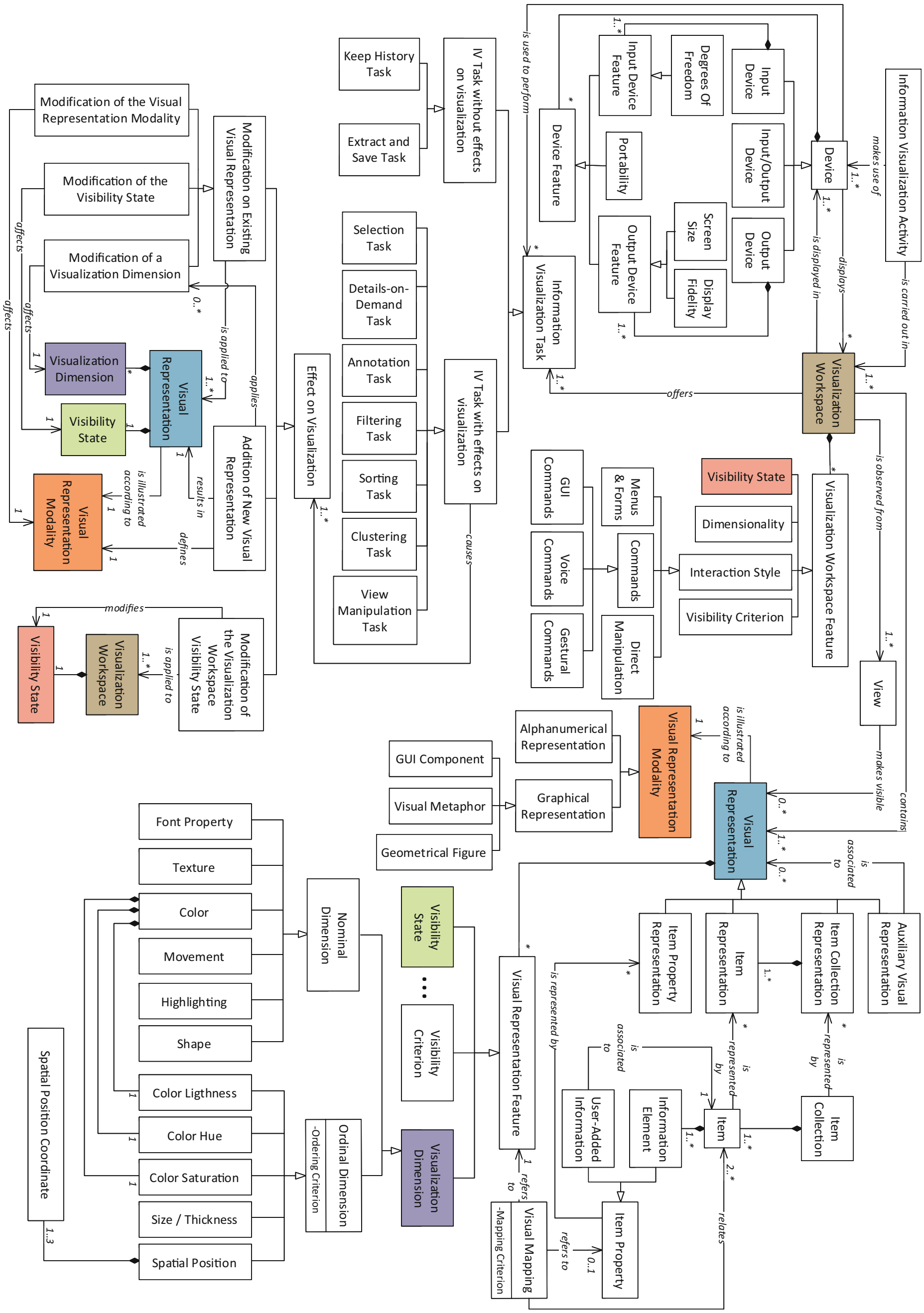


Fig. 1 Information visualization conceptual model

model supports representing information visualization activities that are performed in a multi-device setting, where visualization workspaces are displayed in different devices.

An information visualization activity, for example, could be carried out at the same time in two visualization workspaces, one of them displayed in a tablet and used to represent a corpus of documents in a given way, while the other one could be held in a computer screen and could display the same set of documents but in a different way. In this case, manipulating information in any of the visualization workspaces should cause effects in the other visualization workspace to maintain the consistency. For instance, if a user selects three document titles from the list in one of the visualization workspaces and deletes them, the visual metaphors used to represent these documents in the other visualization workspace should automatically disappear.

A *Visualization Workspace* can be defined as a digital environment where an information visualization activity is performed. One of the interesting aspects of digital environments is that they can be infinite. This is especially relevant for information visualization as usually the amount of information to be displayed is huge.

The problem, then, comes because the screen of the output device has a limited size and then, when the information occupies too much space in the digital environment, it cannot be displayed as a whole, or at least not in a way that it is visible and processable by the user.

To solve this problem, the user should be allowed to navigate through the environment by manipulating the *View* from which he/she observes the visualization workspace so that, at a given moment, not all the visual representations that are present in the visualization workspace would be visible for the user, even if they actually exist in the environment.

In Google Scholar, for example, the visualization of the collection of documents returned as result of a query is performed in a unique visualization workspace. At every moment, the user can only visualize this workspace through a view where only ten or twenty documents, depending on the user configuration, are visible and manipulable. Then, if the user wants to see the rest of the documents, he/she has to change the view.

Alternatively, 3D Explorer offers two workspaces are offered. One of them is a three-dimensional environment where a set of coloured lines, all stemming from the same origin, encode the clusters into which the documents are classified according to the topic they deal with (see Fig. 2).

The second information visualization workspace represents the same documents, also clustered by the topic they address, but as a one-dimensional textual list (see Fig. 3). In this workspace, colour is again used to encode the thematic clustering (see element (a) in Fig. 3) and the colour hue is used to encode that the documents can be categorized in more than one thematic cluster (see element (b) in Fig. 3). Additionally, document representations are sorted based on two criteria. First of all, each document is inserted in a drop-down list that represents the topic (see element (c) in Fig. 3). Secondly, within each of these drop-down lists, the alphanumerical document representations are sorted alphabetically (see element (d) in Fig. 3).

Then, each visualization workspace has to be considered as an independent digital environment that has its own features. Among them, a visualization workspace can be visible or not (this concept is called *Visibility State*).

This decision may be taken by the user but, occasionally, the information visualization system can automatically display, or at least suggest to display, a given visualization workspace depending on the context in which the visualization is being performed, on the information visualization task the user is currently performing or on the user preferences. In the model, this concept is reflected as *Visibility Criterion*.

As an example, both Google Scholar and Calimaco have a unique visualization workspace, and therefore the visibility criterion is also unique: the workspace is always visible. However, in the case of 3D Explorer, the three-dimensional workspace is always visible, while the one-dimensional workspace is only visible on-demand.

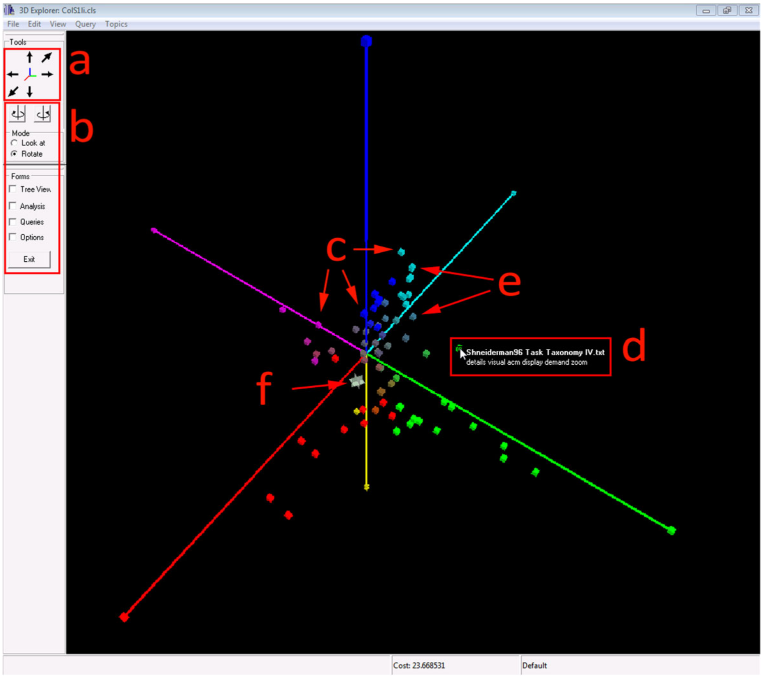


Fig. 2 First workspace of 3D explorer

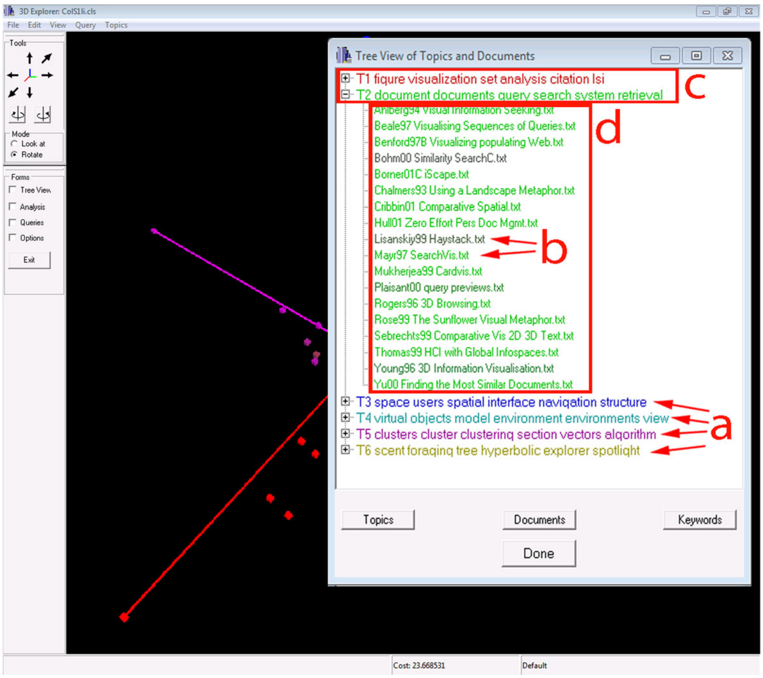


Fig. 3 Second workspace of 3D explorer

There are other features that specifically depend on the aim of the information visualization system and on the specific design of its visualization workspaces, such as the *Interaction Style* it follows, or its *Dimensionality*, which is the number of dimensions it uses to present the information. In our model, we have represented the three interaction styles proposed by Shneiderman et al. [76]:" *Menus, Forms and Dialog Boxes"* used in Google Scholar," *Commands"*, and" *Direct Manipulation"* used in Calimaco and 3D Explorer. However, new ones can be added if required.

With respect to the dimensionality, a visualization workspace can be:

- *One-dimensional*, when information can only be located along one axis. Visualizations that illustrate the information as a list of items, like Google Scholar, are an example of one-dimensional workspaces;
- *Two-dimensional*, if the information is distributed along both horizontal and vertical axes but depth is not considered, like in FilmFinder [3] or in GRIDL [75]; and
- *Three-dimensional*, if information is positioned in a 3D space, like Calimaco or 3D Explorer.

Most likely, there exist many other concepts that could define a visualization workspace. However, the present model is not intended to represent them all, but only those that are considered common to any information visualization system. Nonetheless, it is flexible and expandable enough to be adapted or augmented, if needed.

5.2 Devices

First of all, it is necessary to define which kind of *Devices* are going to be used in the information visualization system, as the designer has not only to consider how the information is going to be visualized, but also how it is going to be manipulated.

There are *Output Devices*, for example a monitor, a projector or a head-mounted display; *Input Devices*, for instance a keyboard, a microphone or an optical tracking system; and *Input/ Output Devices* that allow transmitting information in both ways, like smartphones, tablets, touch tables or tangible devices.

Each of these devices has a set of associated features that allow describing their functionality, capabilities, and restrictions. Our conceptual model illustrates some of the aspects that have to be considered to design how the user visualizes and interacts with the information. Some of these *Device Features* can characterize any device, regardless of its type, while others may apply only to input or output devices:

- *Portability*. It has to be considered if the device can be carried, and then the system can be used under different environmental conditions, or if, on the contrary, the device has a permanent location. In the case of input devices, this can influence how the user interacts with the system because of physical restrictions.
- *Degrees of Freedom*. It is essential to know how many degrees of freedom the device is allowed to manipulate, as this will directly affect which kind of interaction techniques can be performed in the visualization workspace, and then it indirectly also affects the visualization design. As an example, the navigation through a three-dimensional virtual environment is very different when the user uses a device that can manipulate only two degrees-of-freedom, like a classical mouse, as opposed to using a six degrees-of-freedom

device, like a 3D mouse. In the first case, the degrees-of-freedom that are not directly manipulable with the device have to be manipulated through software-based solutions, like 3D widgets allowing to translate and rotate the view around any axis. This aspect is closely related to the dimensionality of the workspace.

- *Display Fidelity*. It is essential to know which is the fidelity level of the device when displaying visual features like colour and texture. In fact, if the definition is not good enough, there is a risk that some of the values of some visual features may not be displayed as expected. This would imply loss or misunderstanding of the information carried by these visual features.

As it happens with the visualization workspace features, the goal is not to exhaustively represent all the features that can define a device, but only to point out some of those that are relevant and should be considered when designing an information visualization system.

5.3 Visual representations

In terms of visualization, one of the first design decisions that has to be taken is how the items of information are going to be displayed. In the literature, there are many works defining all the aspects related to the visual representation of the information, like which are the different types of representations that exist and which are their main characteristics, or when and how to use each of them [8, 9, 18, 23, 27, 28, 44, 61, 64, 67, 68, 78, 91, 92].

In the proposed model, it is considered that *Item Representation* is the core of the visualization, as it conveys information not only about the item's content, but also about its relationships with other items. In the field of scientific research, for example, the information-seeking process mainly revolves around the documents [53].

Besides this, our model also considers the possibility of representing an *Item Collection*. In this case, the collection can be illustrated as a single visual representation, or as an aggregation of item representations. Following the previous example, researchers typically need to manage many very large collections of documents, either from own or external sources, and then they may find useful to represent the collections themselves instead of the documents. In the first case, a collection of documents that are represented through a sphere could be itself represented as a new sphere that would conceptually and visually include all the documents belonging to it. In this case, for example, the size of the spheres could represent the number of documents that belong to the collection. In the second case, the organization in the visualization workspace of the items can indirectly illustrate a collection. This, for example, can be done by using thin lines that separate groups of items (as it is done, for example, in a Scatterplot Matrix), but also by using empty spaces to visually cluster the data (based on the proximity Gestalt's principle).

Our model also takes into consideration that, on occasion, users may need to visualize more specifically some of the properties of an item, instead of the item itself. This is especially relevant when the item is a complex entity that a lot of heterogeneous data, such as a scientific document that contains information like its authors, publication venue and date, keywords, etc. In these cases, each of these data can be visually instantiated as an *Item Property Representation*.

However, beyond the main information items and their properties, there are other aspects that need to be represented in a visualization workspace. These *Auxiliary Visual Representations* are needed to represent, for example, the control widgets needed to execute a task.

For instance, Google Scholar provides a set of filtering, sorting and configuration options through selectable textual elements, check-boxes, buttons and visual metaphors (see elements (c) and (f) to (j) in Fig. 4), while Calimaco offers the user eight visual metaphors to perform different information visualization tasks (see element (a) in Fig. 5).

3D Explorer, meanwhile, offers two types of control widgets: the first one allows rotating the view around the vertical and horizontal axes, and to zoom in and out by the use of a set of arrows (see element (a) in Fig. 2). Additionally, it also allows to rotate the view using the camera as reference point (instead of the axes), by selecting the *"Look at"* mode in a component of the GUI. This component makes part of a set of radio-buttons, check-boxes and buttons (see element (b) in Fig. 2) that are provided in order to launch some functionalities, like changing the rotation mode, making the view rotate automatically, making visible the second visualization workspace, called" *Tree view"*, modifying the number of topics into which the documents are clustered, representing a query in the information space as a new virtual document, or modifying the view options.

An auxiliary visual representation can also be used to encode the legend of the visualization. A legend is especially useful when the workspace makes use of a variety of visual features to convey different kinds of information about the items, or to explain to the user the meaning of some graphical representations, especially if they are abstract. For instance, the system could use a star as a metaphor to encode that a document has been marked as favourite by the user, or it could also draw a pin near the document that has been blocked by the user in order to avoid it being affected by filtering tasks. In both cases, it would be probably needed to explain them in the legend. It could also be possible to make use of a label, which is visually represented through an alphanumerical representation. For instance, if colour is used to encode the publication date of a set of documents, a label with the text *"Before 2000"* written in red, and another label with the text *"After 2000"* written in green would perfectly explain the use of the colour in the visualization.

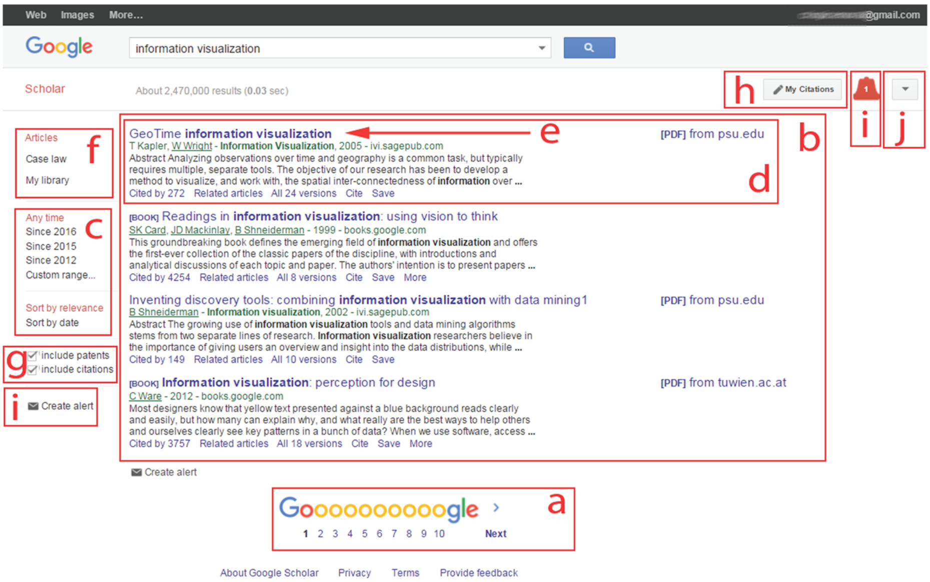


Fig. 4 Google Scholar

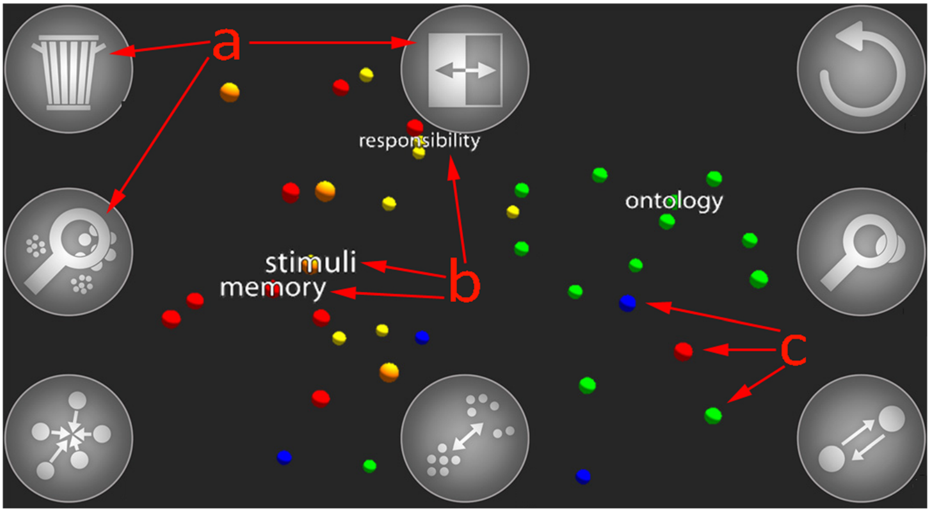


Fig. 5 Calimaco

It may be also needed or desired to illustrate in an information visualization system some labels associated to other visual representations in order to provide some kind of information about the latter. The visual representation of a document collection, for instance, can have a label attached (that is, an auxiliary visual representation) in order to indicate the number of documents that are part of the collection. A document can also have a label with the most frequent term appearing in the document, as it is done in Calimaco (see element (b) in Fig. 5). Additionally, a control widget can have a label indicating the name of the functionality or information visualization task it triggers, as it is done in 3D Explorer (see element (b) in Fig. 2).

Auxiliary visual representation can be also instantiated to encode elements related to the interaction. In Calimaco, for instance, a set of connecting lines link the cursor (see element (b) in Fig. 6), represented by two fingerprints, to all the document representations that are close to it (see elements (a) in Fig. 6). This visual representation is extremely useful as the workspace is three-dimensional, and then it may be difficult for the user to determine how close are the fingerprints to the document representations that he/she wants to interact with (especially if the output device is two-dimensional). Additionally, if a document representation is selectable, then the line connecting it to the cursor becomes bold and dashed (see element (c) in Fig. 6).

Lastly, relationships between items (or item properties) can also be encoded through a set of auxiliary visual representations, like a line connecting two related item representations, or a square enclosing a set of related item representations. This is indirectly done in Google Scholar, where a transparent box surrounds all the properties related to a document (see element (d) in Fig. 4).

Regardless of the concept they represent, visual representations have a set of features that are related to their visibility, and to their visual appearance. Like visualization workspaces, each visual representation has a *Visibility Criterion* and a *Visibility State*.

Finally, there are different types of data that can be illustrated, requiring the use of different visual representations. Shneiderman [74] classifies the data that can be visualized in an information visualization system into seven categories: *One-*, *Two-*, *Three-* and *Multi-dimensional Data*, *Temporal Data*, and *Tree and Network Data*. Information-seeking performed by

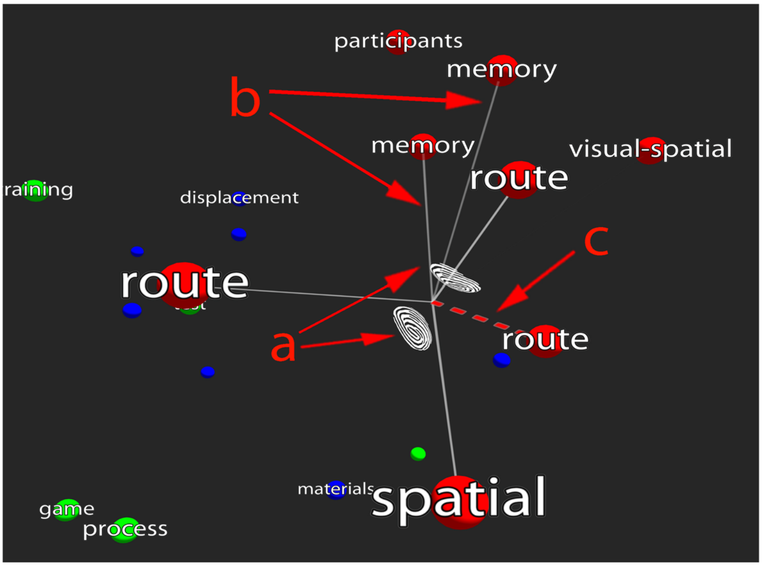


Fig. 6 Example of use of other visual representations in Calimaco

researchers, for example, deals with multi-dimensional and network data, as documents and document collections aggregate a set of data, and many of these data allow linking documents and document collections [51].

In the proposed model, we have categorized all the possible ways to encode a visual representation (called *Visual Representation Modalities* in our model) into two main groups: *Alphanumerical* and *Graphical Representations*.

Alphanumerical representations cover the use of any kind of textual representation to transmit the information, like the one used in Google Scholar to represent the documents, or the labels used to illustrate the most frequent term of a document in Calimaco (see element (b) in Fig. 5) and of a topic in 3D Explorer (see element (d) in Fig. 2).

Graphical representations, on the contrary, group all the geometrical, pictorial, metaphorical, and non-textual representations. Among them are, for example, *GUI Components*, *Visual Metaphors*, and *Geometrical Figures*. GUI components include the interaction elements that are typically used in software systems, like menus, buttons, or sliders. Visual metaphors encode, more or less reliably, real concepts, like a star, a drawing pin, or the bin icon. Finally, geometrical figures refer to any geometrical representation that conveys information. For instance, a circle or a square that encloses a group of item representations, or a line linking two item representations, probably reflects that they are somehow related (like the invisible box surrounding the properties of a document in Google Scholar).

5.4 Visualization dimensions

Each visual representation has associated a certain number of visual features (called *Visualization Dimensions* in our model) that can be modified in order to provide some kind of information. Conceptualizing these dimensions is essential, as the information to be visualized is often very complex, and then it requires using more than one aspect of the visualization.

This aspect is especially relevant as there are several visual dimensions that can be used in an information visualization system, and each of them inherently provides different kinds of information. In fact, in the literature, visual dimensions have been divided into four main groups [17].

First of all, there are features related to the use of the space to encode information, like position and orientation. These features have a great importance in information visualization as humans are especially good at processing spatial information. That is why spatial location is commonly used to encode relationships between items, taking advantage of a set of organization principles, called *"Gestalt Principles"*, that affect visual perception [79]. As an example, two objects that are very close to each other are intuitively considered more similar or related by the observer than two objects that are distant from each other [11].

There is a second group of properties that are called retinal, as the eye's retina is directly sensitive to them. These visual properties are used either to encode some characteristics of the data, or to reflect how these data are related. Using these kind of properties is a key factor in information visualization as it allows to perform an automatic processing, which is *"superficial, parallel, can be processed non-foveally, has high capacity, is fast, cannot be inhibited, is independent of load, unconscious, and characterized by targets 'popping out' during search"* [19]. This group includes visual features like colour and size.

There is a third type of features that allow identifying relationships between data: connection and enclosure. Connection refers to the use of an explicit visual representation to relate two elements, as for example a line. Enclosure occurs when a set of elements create a visual set that is delimited, for example, by an explicit border, or by an invisible outline of white space.

Finally, when animation is used in the visualization, the variation of the visual representation along time also allows encoding information about the data. As an example, the rotation speed of a visual representation can provide some information to the observer.

All these visual features have a main characteristic in common: they are preattentive, and then they ensure a fast and efficient impact on the user, but without cognitively overloading him/her.

In our model, these features have been categorized into two groups depending on whether their values can be sorted based on an *Ordering Criterion* (*Ordinal Dimensions*), or not (*Nominal Dimensions*).

Some examples of ordinal dimensions are:

- *Spatial Position*. Within the visualization workspace, a visual representation can be located in positions determined by one to three spatial position coordinates (axes x, y, and/or z), depending on the dimensionality of the visualization workspace. Focusing on each of these three axes, a visual representation can be above or below other visual representations, to the right or to the left, or closer or farther. For example, in Google Scholar, documents are presented in an ordered list based on their relevance with respect to the query (see element (b) in Fig. 4). Then, the vertical spatial position coordinate of a document representation depends on its relevance. In Calimaco, on the contrary, document representations are located in a 3D space, according to their topic, so that documents dealing with similar or related topics are represented in close positions in the visualization workspace. Similarly, in 3D Explorer the location of the document representations depends on the topics they deal with, but in this case the spatial position is figured out with respect to the lines that represent the different clusters. Then, the closer a document is represented to a given cluster line, the more its content is related with the topics represented by the cluster.
- *Colour*. In its basic form, colours can be used as a set of unordered and unrelated values. As an example, in Google Scholar, documents' properties are encoded using text entries.

Each of these properties provide a different type of information (see element (d) in Fig. 4) and is displayed with a different colour in order to facilitate their visual recognition by the users: the title is written in blue, the authors in green, and the abstract in black. In the case of Calimaco and 3D Explorer, colour is used to represent the document's degree of membership to clusters (see elements (c) in Fig. 2 and Fig. 5). In addition, colour admits different levels of detail, and depending on how it is used and manipulated, it can be also considered as an ordinal dimension. For this, it has to be divided into a set of ordinal dimensions, like those proposed by Joblove & Greenberg [36] in the HSL model, where colours are decomposed into 3 aspects whose values can inherently be sorted: *Colour Saturation*, *Colour Hue* and *Colour Lightness*.

- *Size or Thickness*. A visual representation can be bigger/thicker or smaller/thinner than other visual representations. As an example, in Google Scholar the size of the fonts used in each of the document properties (title, authors, abstract...) is different (see element (e) in Fig. 4). In Calimaco and 3D Explorer, depending on the level of zoom of the view, document representations are bigger if the camera has zoomed in, or smaller if the camera has zoomed out.

On the other side, some of the nominal dimensions that can be used to encode a visual representation, besides the colour in its basic form, are:

- *Shape*. A visual representation can be, for example, a circle, a sphere, a square, a triangle, or a hexagon, but also a house, a drawing pin, or an arrow.
- *Texture*. As background, a visual representation can have some kind of pattern, like parallel lines that can be drawn vertically, horizontally, or in diagonal; grids; or regularly located geometric figures, like points, squares, diamonds, etc.
- *Highlighting*. Each visual representation can have associated a binary state: highlighted or not. For example, a bright light surrounding a visual representation can be used to distinguish the representations that have been selected from those that have not.
- *Movement.* At a basic level, a visual representation can rotate, jump, or vibrate, for example, to encode a binary value, as done with highlighting. Additionally, if the movement speed can be modified, it can reflect multivalued states. In this case, speed can even be considered an ordinal visualization dimension.
- *Font Properties*. In the case of alphanumerical representations, there are a number of properties related to the text font that can be modified, like the type of font, for example Arial or Times New Roman; or its format, that is serif or sans serif, bold, underlined, italics, etc. Font size is not included in this category, as its values can be sorted, and then it is an ordinal dimension.

Many of these visualization dimensions have been already identified as retinal properties that facilitate the preattentive processing of what the user is seeing. Among the visualization dimensions proposed in the literature that are considered preattentive, are spatial position, colour (both in its basic form and decomposing it), texture, shape, size, direction of motion, and flickering [10, 17, 31, 42, 43].

The specific type of visualization dimension to be used depends on the kind of information that has to be displayed and on the aim of the visualization. There are three types of data that can be displayed in an information visualization system: nominal, ordered and quantitative [17].

If the aim is, for example, to cluster a set of document representations according to a given criterion, then using either a nominal dimension or an ordinal dimension is equally possible. In this case, the decision of which one to use should be based on the characteristics and preferences of the user, so that the selected visualization dimension is as salient as possible for the user.

However, if the information to be illustrated in the visualization expresses some kind of order, or can be sorted according to a certain criterion, then using an ordinal dimension is required. For example, if document representations have been clustered according to their publication date, it could be interesting to visualize the created clusters in a chronological way (which is an ordering criterion), so that the representation of the newest documents would be located to the right of those published in previous years. Analogously, the colour saturation of the document representations could be used to encode the same information. Then, the oldest documents would be represented with a poorly saturated colour, and the saturation would increase as the documents are newer.

5.5 Information visualization tasks

One of the most cited and used guidelines when designing information visualization systems is the mantra proposed by Shneiderman [74], where he identifies seven functionalities that should be present in any information visualization system, regardless of its domain, its graphical design and the type of the data displayed: *"Overview"*, *"Zoom"*, *"Filter"*, *"Details-on-demand"*, *"Relate"*, *"History"*, and *"Extract"*.

In our model, the concept *Information Visualization Task* is defined in order to cover all the functionalities that an information visualization system may offer. In our case, however, we have distinguished between the tasks that result in a certain modification in the visualization (*Information Visualization Task with effects on visualization*) and those that do not affect the visualization (*Information Visualization Task without effects on visualization*). Although our model is more focused on the first type of tasks since it focuses specifically on the visualization, we have also included the second one so that the model is as complete as possible.

The first group includes most of the tasks enunciated by Shneiderman in his mantra. Both *"Overview"* and *"Zoom"* functionalities are covered in our model by the *View Manipulation Task* that includes all the subtasks that imply modifying how the visualization workspace and its content, the visual representations, are shown to the user. Then, his *"Filter"* functionality is directly mapped in our model into the *Filtering Task*, where the use of control widgets indicates which conditions must be met by the items in order to be visible. Shneiderman's *"Relate"* functionality is also partly represented in our model by the *Clustering Task*, which is one of the subtasks allowing to relate items. Finally, the *"Details-on-demand"* functionality proposed by Shneiderman is directly reflected in our model, using the same name.

The second group can contain, among others, the two remaining functionalities suggested by Shneiderman: *"Keep history"* of actions to support undo, replay, and progressive refinement and *"Extract and Save"* subcollections and query parameters. In both cases, the execution of the tasks does not translate into an effect on the visualization, but they are very relevant to facilitate the use of the information system. In fact, saving each of the actions performed by the user in the system is useful, but not only to improve its usability, but also to be able to model the user, his/her preferences, and his/her intentions so that the system can adapt to his/her specific needs at every moment.

In addition to these tasks, in this work we propose to consider three more tasks with effects on the visualization when designing an information visualization system. First, as in many cases the data included in the information items can be sorted, special attention has to be paid to the *Sorting Tasks* that allow organizing a set of information items based on a given ordering criterion.

Additionally, one of the most repetitive and unsupported tasks (or supported, but in a non-usable way) while seeking for information is the *Annotation Task*. In our opinion, this task has to be considered in the design of most visualization systems, as annotating an information item does not only means writing a series of notes or summaries, but also assigning tags, indicating that the item is a favourite, or underlining relevant content. All these actions have a direct consequence on the visualization, and their proper representation can be used to facilitate the classification, recognition and processing of the notes or comments.

Finally, a *Selection Task* is included even if it is more a means than a visualization task, as it is mainly used as a previous step required to perform other tasks like obtaining details-on-demand or annotating an item. This task is especially important as it has a great influence in the interaction with the visualization and allows the user to understand which set of representations he/she is working with at a given moment. Additionally, in a context where hundreds or thousands of items have to be represented at the same time, it is not obvious how to design one or more selection techniques that allow the user to select the item(s) he/she wants in an accurate and efficient way.

Calimaco, for example, supports five of these information visualization tasks, as the user can:

1. Select documents' representations (selection task).

   This is done by pinching the index and thumb fingers when the fingerprints (represented by the element (a) in Fig. 6) are close enough to a document that is selectable (this is illustrated through a dotted and coloured line, as shown in element (c) in Fig. 6);
2. Obtain details-on-demand (details-on-demand task).

   As an example, to display the content of the document, the user has to manipulate the 'inspect' widget (see middle right element (a) in Fig. 5) once the targeted document has been selected;
3. Remove representations from the visualization (filtering task).

   By using the 'bin' widget (see upper left element (a) in Fig. 5), the user can remove from the visualization the documents' representation that have been previously selected;
4. Classify documents based on the topics they deal with (clustering task).

   This is done by manipulating the 'clustering' widget (see lower central element (a) in Fig. 5), that allows indicating in how many thematic groups the user wants to categorize the set of documents being displayed;
5. Change the view of the workspace (view manipulation task).

   For example, the user can manipulate the 'jump' widget (see middle left element (a) in Fig. 5) to modify the viewpoint of the camera (each 'jump' centres the view on one of the groups of the clustering). The user can also translate the view of the workspace by moving the fingers in the open position (thumb and index are perpendicular) in any of the three axes (translation in the z axis corresponds to zooming in and out). The user can also rotate the view by rotating his/her fingers in the closed position (thumb and index are pinched) around any of the three axes.

As reflected in Fig. 4, Google Scholar also provides four of these tasks:

1. Get more information about a result (details-on-demand task).
   This task is performed, for example, by clicking in the document's title (see element (e) in Fig. 4);
2. Filter the obtained results (filtering task).
   The user can filter the results, for example, based on the publication date of the document (see element (c) in Fig. 4) or on the type of document (see element (g) in Fig. 4);
3. Manipulate the view (view manipulation task).
   The user can move to a previous or next page to see more results by clicking on the numbers, arrows or *"Previous"* and *"Next"* labels (see element (a) in Fig. 4)
4. Order the results according to a given criterion (sorting task).
   The user can sort the result based on the relevance or on their publication date (see element (c) in Fig. 4).

Finally, 3D Explorer allows performing four tasks:

1. Select a document representation (selection task).
   The user can select a document representation just by clicking over its representation (see element (c) in Fig. 2);
2. Cluster the documents according to their content (clustering task).
   The user can indicate how many clusters he/she wants (through the *"Topics"* item of the menu, see Fig. 2), and the system classifies them accordingly (each cluster is represented through an axis and a colour);
3. Obtain details-on-demand (details-on-demand task).
   Once a document representation is selected, its filename and keywords are displayed (see element (d) in Fig. 2). In addition, the user can read the content of the document by double-clicking over its representation. In the Tree Explore view (see Fig. 3), the user can also click on the clusters' representation (see element (c) in Fig. 3) to show the related documents or click on a document's representation (see element (b) in Fig. 3) to display its keywords;
4. Manipulation of the view (view manipulation task).
   As it has been explained before, the user can translate, zoom in and out, and rotate the view by manipulating a set of widgets (see elements (a) and (b) in Fig. 2).

Nonetheless, the information visualization model is flexible and scalable enough to admit the addition of new concepts, such as new information visualization tasks. As an example, the information visualization tasks implemented in Calimaco that modify the spatial position of the document representations by modifying the inter-document and/or intra-cluster attraction forces could be perfectly added to the model. This sorting task is performed by manipulating the intra-cluster force and inter-documents force widgets (see lower left and right elements (a) in Fig. 5).

5.6 Effects on visualization

Despite not all the information visualization tasks that may exist in an information visualization system have been considered, it can be asserted that these tasks can mainly produce three kinds of *Effects on the visualization*:

1. *Addition of a new visual representation to the visualization workspace*. For instance, a new graphical representation, like a container, can be drawn in the visualization workspace in order to encode the result of a clustering task. As another example, 3D Explorer allows representing a specific query as a new document whose representation is located in the workspace according to its thematic similarity with the documents. This task, even if it is not explicitly considered in the proposed model, implies the creation of a new visual representation to illustrate the query (see element (f) in Fig. 2).
2. Some of the visual representations are somehow affected by the execution of the task:

   a) *Modification of one of the Visualization Dimensions of some visual representations*. For instance, in Calimaco and 3D Explorer, after performing a clustering task, the colour of document representations changes to encode the membership to the clusters.
   b) *Modification of the Modality* of some visual representations. A clear example of this is the representation of the documents and folders in the file explorer of an operating system, like Windows Explorer or Apple Finder. In these environments, which actually act as information visualization systems, the elements can be represented either by icons (that is, visual metaphors) or by textual entries (that is, alphanumerical representations), depending on the visualization mode selected by the user.
   c) *Modification of the Visibility State* of a set of visual representations. For instance, the user may want to visualize only those documents he/she has marked as favourites, and then the rest of the document representations have to become invisible. As an example, in Calimaco the user is allowed to select and remove from the visualization a single document representation, a set of them selected manually, or all those that belong to the same cluster.

3. *Modification of the Visibility State of a Visualization Workspace*. As an example, if a details-on-demand task is performed, a new visualization workspace displaying the content of the selected document can appear in the visualization. This occurs in 3D Explorer, where the second workspace can be visible or not depending on whether the user activates it or not.

5.7 Visual mapping

One of the most useful feedbacks that can be obtained by visualizing a set of items in an information visualization system consists in identifying how these items relate to each other. In cases where items contain a lot of data that can be used to establish relationships between them, this functionality becomes essential.

This is the reason why the model includes aspects that have to be considered to visually encode the relationships existing between items. First of all, it is essential to know which kind of data is going to be used to determine if two or more items are related. For example, in the case of document collections, the relationship can be established according to any of the Item Properties, implicit or explicit, of a document. Properties like the authors names, the topic, the publication date, or the publication venue are represented in the concept *Information Element*, whereas *User-Added Information* includes any information the user may add, like personal notes, tags, highlighted terms or favourite documents [51].

Secondly, the designer has to specify the *Visual Mapping* that defines how the relationship between items is illustrated in the visualization system. For that purpose, the designer has to detail the *Mapping Criteria* that indicates how the features of the related visual representations have to be modified to illustrate the relationships based on the items' properties. For example, Calimaco makes use of the colour to represent the clustering of the documents based on their content. Then, the mapping criterion could be, for example, that all item representations belonging to cluster 1 are red, those belonging to cluster 2 are green, and those included in cluster 3 are represented in blue. Another criterion could be that each new cluster is coloured with a new randomly assigned colour.

The choice of what features of the visual representations should be used for relating items has be carefully analysed in order to ensure that the mapping can be done properly. It is essential to ensure that the selected visual representation feature has enough distinguishable values to encode all the different possible values of the item property to be displayed. For instance, if the user wants the system to illustrate how documents are related based on their publication date, using the size property would probably not be the best choice, as it would imply having as many distinguishable sizes as publication years, which could be dozens. In this case, for example, the spatial position along the horizontal axis could be used to chronologically locate the documents, or to draw containers that enclose the representations of the documents published in the same range of years.

6 Application of the model

The proposed model can be used for three purposes: to describe, compare, and as guideline to design information visualization systems.

6.1 Description

First, it can be used as a framework for characterizing existing information visualization systems. For this purpose, we believe that the instantiation of our diagrams to describe an existing visualization application helps to perform a complete inspection of the latter, as the model can be used as a kind of checklist. This has been done above, as we have thoroughly described three information visualization systems (through the examples provided), by detailing one by one all the concepts and relationships that intervene in them.

Throughout the paper, we have used the model to describe an information visualization system allowing researchers to seek for information. However, as a general model, it could be also used to describe information visualization systems or elements displaying any kind of data, like for example the map of a metro network or any of the well-known techniques used to display hierarchically organized information, like SemNet [28], which is used to display knowledge databases like ontologies through a graph-based representation, or Starlight [67], which represents 3D objects in the space. In both cases, the proposed model enables to fully describe them through:

1. the visualization dimensions that are used, like colour, shape, size, or position;
2. the items that are represented through different representation modalities. SemNet uses geometrical representations, while Starlight also uses metaphors, like maps;
3. the tasks affecting the visualization that are provided, like changing the position of an item or groups of items, or modifying the point of view of the full workspace;
4. the relationships between items that are represented through other representations, like lines, i.e. connectors, or by modifying some visualization dimensions of the items, like the colour.

In order to validate and demonstrate the model's flexibility and generality, we have partly modelled different examples of information visualization made with the general purpose library D3 [12]. Fig. 7, for instance, shows a stereographic projection of many of the countries of the World (and the Antarctica) where two data are provided through the visualization: size illustrates the area of the country, and colour reflects the continent to which it belongs. In Fig. 8, we describe this information visualization example using our model.

We also present as example of visualization made with D3 a SPLOM representing the well know dataset of the iris flowers on the Gaspé Peninsula [6]. In this visualization, illustrated in Fig. 9, each scatterplot can be considered as a collection of items. In Fig. 10, for instance, we show how the model describes the representation of one of the items (the Virginica flower with the lowest sepal length).

Evaluating the information visualization designs and/or systems through the assessment of the value and benefits these visualizations provide is also essential [80, 84]. As stated by Behrisch et al. [7], describing in detail an information visualization system is essential to assess its quality. In their work, they define 3 levels of metrics that must be used to assess the quality of an information visualization system: low-level, mid-level and high-level metrics dealing, respectively, with perception, interaction tasks and user cognition. Our model provides a large coverage of the two first levels, as it allows describing the visual variables and features used in the system (low-level metrics), and the interaction tasks used to manipulate the system (mid-level metrics).

6.2 Comparison

Our model can be also used as a means to compare the features and capabilities of different information visualization systems. For instance, Fig. 11 and Fig. 12 are used to compare the

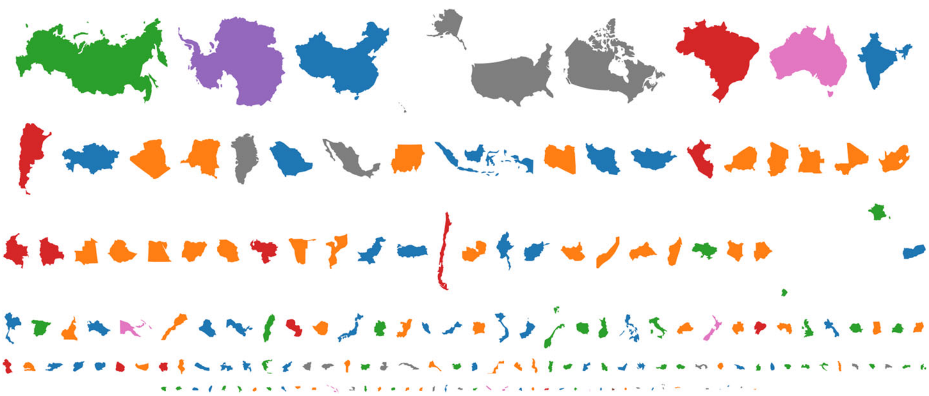


Fig. 7 Countries and Antarctica projected using the conformal stereographic projection, rescaled and sorted by land area (Extracted from [23])

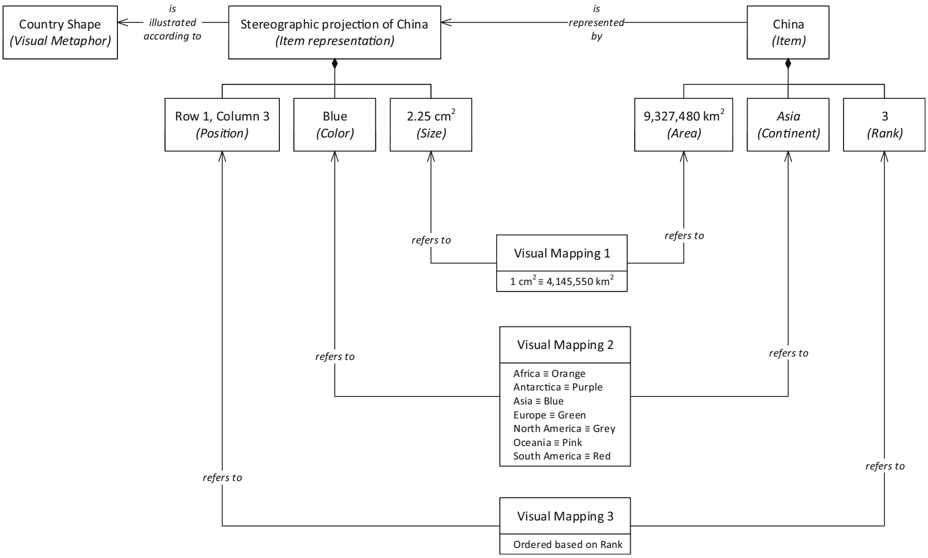


Fig. 8 Description of the information visualization example made with D3 illustrated in Fig. 7

tasks that are respectively available in Google Scholar and Calimaco, and the main effects they cause on the visualization. In both figures, a number is assigned to each of the tasks that are implemented in the corresponding system, whereas those that are not implemented remain in white colour with the text in grey. Then, these numbers are used again to encode which effects on the visualization cause each of these tasks.

Thanks to the visual model, it is very easy and direct to compare different information visualization systems. In the previous case, it stands out that Calimaco allows performing more information visualization tasks than Google Scholar, and that almost all the possible effects included in the model can occur in Calimaco, while Google Scholars tasks can only cause the modification of the visibility state or of a visualization dimension.

All the previous examples, even if they only show a part of the model, clearly illustrate its potential to visually describe and compare two information visualization systems in terms of concepts and relationships.

6.3 Design guidelines

Finally, the model can be also used to guide the design of a new information visualization system by providing a common reference vocabulary, a kind of checklist to decide which tasks and effects should be designed, and a clear framework to map the domain concepts into visualization concepts. As an example, if a new system had to be designed to visualize the collection of music songs of a user, the designer could use the model as a guideline to define several visualization-related aspects such as:

1. how many workspaces have to be designed, and which kind of information should contain each of them;

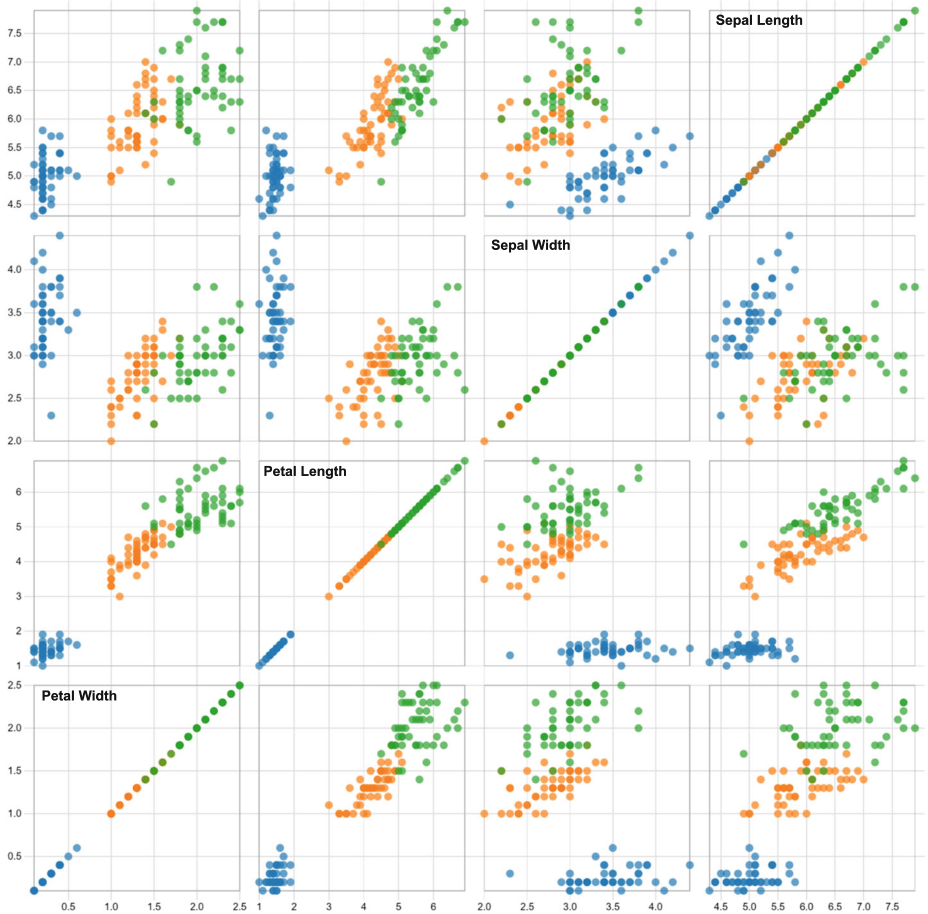


Fig. 9 Pairwise correlations between Sepal Length, Sepal Width, Petal Length, and Petal Width for three species: Virginica (Green), Versicolor (Orange) and Setosa (Blue) (using D3)

2. the characteristics of the elements that visually represent each song, like their modality, shape, colour, or size;
3. the tasks that affect the visualization, and in which way they affect it;
4. the information elements of the songs, like their style, their singer(s), or the year of their publication, that can be used to relate them to each other, and how these relationships are going to be encoded;
5. which of these information elements have to be displayed, and the visual representations of visualization dimensions that are going to be used for that purpose.

7 Comparison with previous models

Much of the research on information visualization has focused on describing the process of converting a set of raw data into an understandable, analysable, and usable visualization. Additionally, the visual features, characteristics and variables that need to be used to achieve

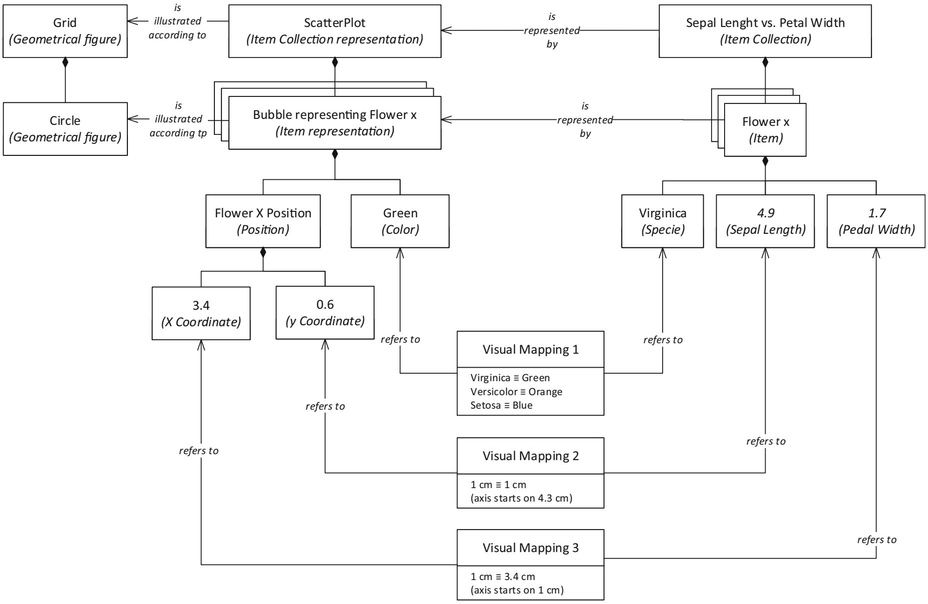


Fig. 10 Description of the information visualization example made with D3 illustrated in Fig. 9

these goals have been extensively defined. However, to the best of our knowledge, none of the previous works have provided a holistic and visual model of the concepts and relationships that need to be considered in an information visualization environment.

Below we compare the proposed model with some of the previous models existing in the literature. As one of the main strengths of our model is that it can be fully represented visually,

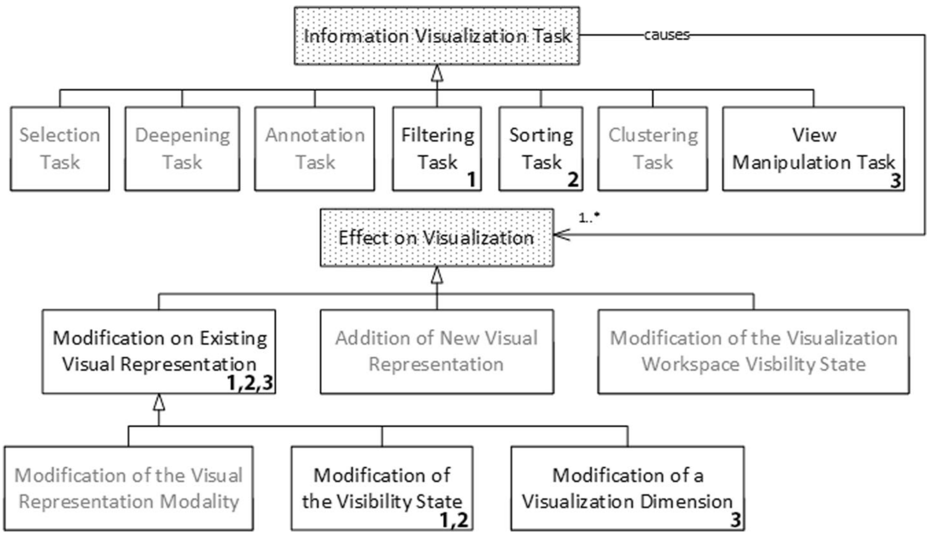


Fig. 11 Description of the information visualization tasks supported by Google Scholar using the conceptual model

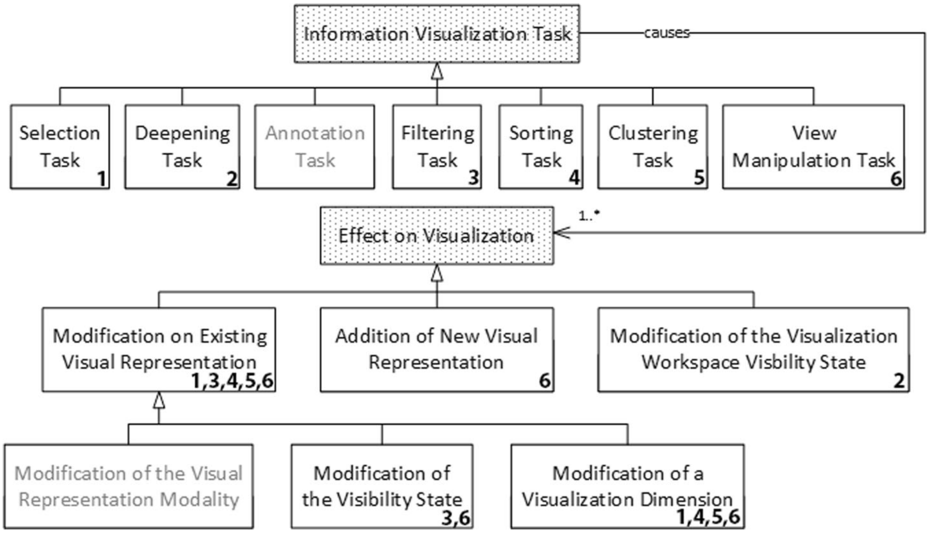


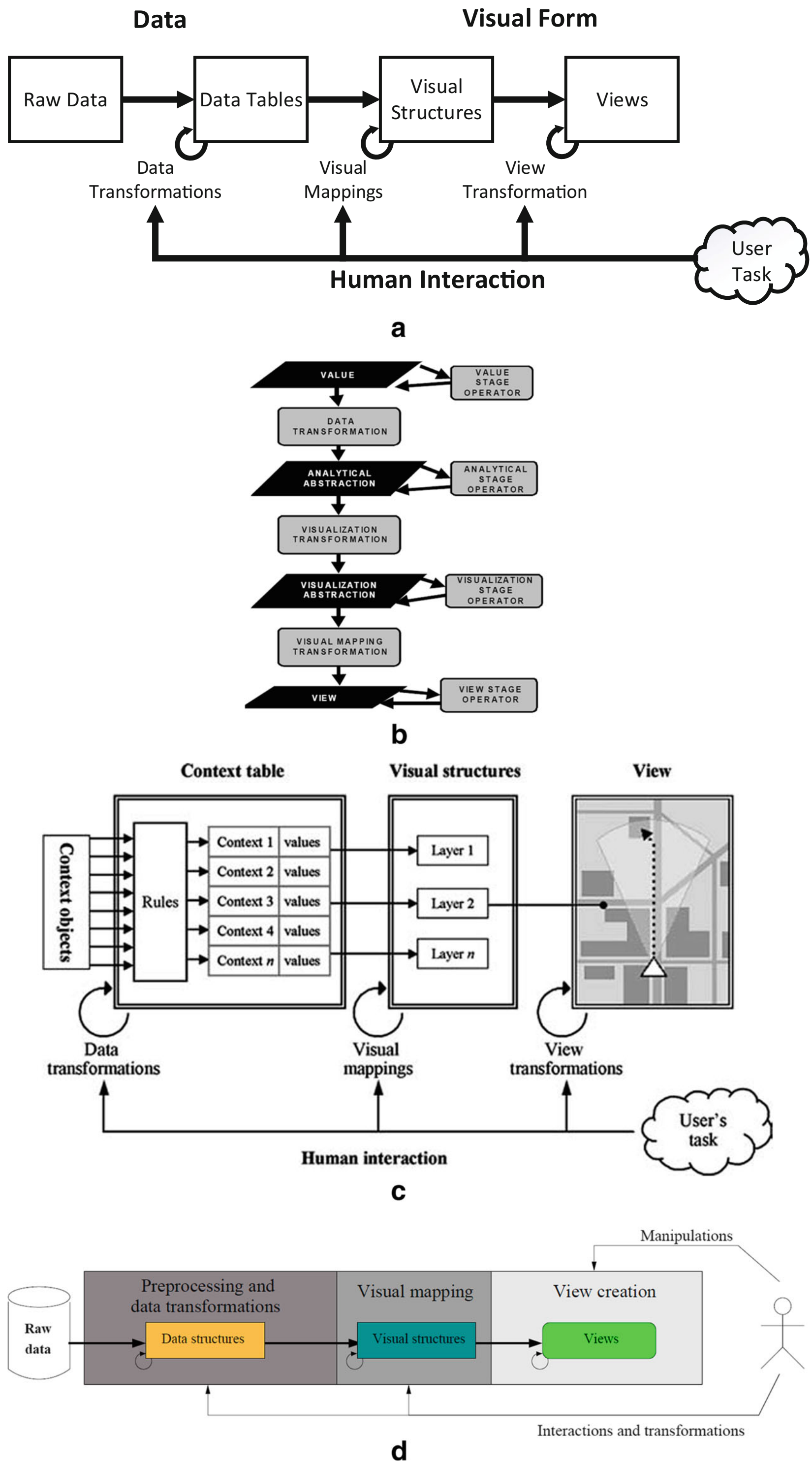
Fig. 12 Description of the information visualization tasks supported by Calimaco using the conceptual model

we compare it with all the existing models and frameworks that are, partially or totally, visually represented.

As illustrated in Fig. 13, there are some differences between the models and frameworks proposed in the literature, but they all agree in their purpose: describing the stages of the process, together with the operations allow to move from one stage to another. In all cases, authors divide the process into a set of stages including the following elements: raw data; processed and structured data; visual structures; and views. All these stages are achieved by performing similar operations: obtaining of the raw data; converting the raw data into structured and homogenised data; mapping the processed data into visual structures; and displaying these visual structures. In all cases, authors highlight the need to consider user interaction, both in the mentioned phases, as well as in the handling of the final visualization of the information. Some of them, also highlight the relevance of defining the domain of the data to be visualized.

In all cases, the first stages of the process do not directly relate to the visualization of the information, but with the cleaning, organization, and structuring of the data. This step es essential to map the data into visual structures. However, in our model we do not contemplate this aspect, as we have decided to specifically focus on the visualization aspects. Additionally, these models present the procedural aspects of the information visualization process, whereas we focus on the conceptualization of the visualization process. Even if the existing models and our proposal are not comparable as they pursue different objectives, they are absolutely complementary. In fact, our model overlaps with the last steps of the presented models, related with the information mapping and the view of the result. The proposed model can be very useful in the visual mapping stage, as it allows the designer to select the best visualization variables and features to illustrate the structured data. Besides, having a single model where all the visual design concepts and relationships are captured together helps the designer to assess and validate his/her decisions before starting the implementation.

Apart from this, in the view phase, the proposed model is useful for modelling the effects of the interaction performed by the user while navigating and manipulating the visualized

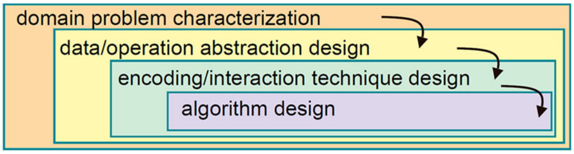


Data
Visual Form
Raw Data
Data Tables
Visual Structures
Views
Data Transformations
Visual Mappings
View Transformation
User Task
Human Interaction
a
VALUE
VALUE STAGE OPERATOR
DATA TRANSFORMATION
ANALYTICAL ABSTRACTION
ANALYTICAL STAGE OPERATOR
VISUALIZATION TRANSFORMATION
VISUALIZATION ABSTRACTION
VISUALIZATION STAGE OPERATOR
VISUAL MAPPING TRANSFORMATION
VIEW
VIEW STAGE OPERATOR
b
Context table
Visual structures
View
Context objects
Rules
Context 1 values
Context 2 values
Context 3 values
Context 4 values
Context n values
Layer 1
Layer 2
Layer n
Data transformations
Visual mappings
View transformations
User's task
Human interaction
c
Manipulations
Raw data
Preprocessing and data transformations
Visual mapping
View creation
Data structures
Visual structures
Views
Interactions and transformations
d

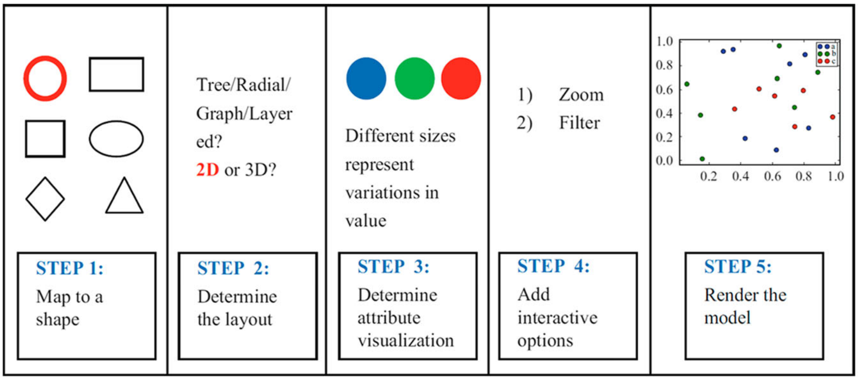
◀ **Fig. 13** Process models of the Information Visualization process. **a** Reference model for visualization (Card et al. [19]), **b** Information Visualization Data State Reference Model (Chi [21]), **c** The reference model for context visualizations (Aaltonen & Lehikoinen [1]), **d** The process of generating a graphical representation (Mazza [45]), **e** Nested design model of visualization creation (Munzner [55]), **f** Steps in the visualization process (Fernandez & Fetais [29]), **g** Adapted Information Visualization Reference Model for immersive Augmented Reality (Irshad [33])

information. This is especially relevant as it allows identifying, in the design phase, which visualization variables and characteristics need to be included to allow performing the desired visualization tasks. As a consequence, the designer is able to ensure that the design of the future information visualization system is consistent and homogeneous.

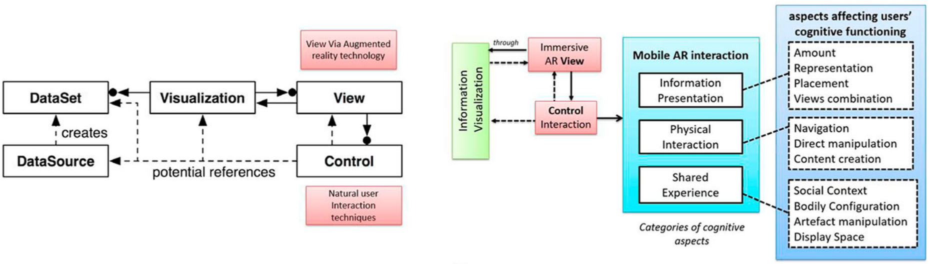
To the best of our knowledge, only Bugajska [16] has proposed a visual non-procedural framework to define the visualization of information. As reflected in his model, depicted in Fig. 14, Bugajska bases her framework in 5 main concepts: *User Tasks*; *Designer Goals*;



e



f



g

Fig. 13 (continued)

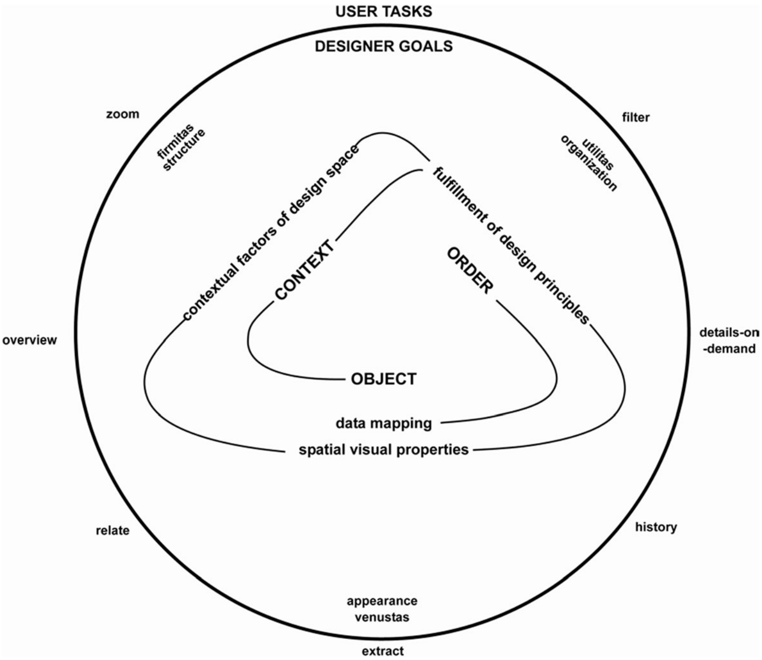


Fig. 14 Framework for the spatial visualization of abstract information (Bugajska [16])

Context; *Order*; and *Object*. As in our case, this framework focuses specifically in the visualization stages of the process, as it deals with the mapping of structured data to visual structures, the selection of visual properties to illustrate the information, and the tasks that may be performed by the user to navigate through the visualized information. However, the pictorial model remains of a very high-level, as opposed to our model, which is self-contained and provides all the details of the visualization.

In all the models presented, it can be observed that there is a lack of a complete definition of all the concepts and relationships involved in the information visualization process. This problem has been partly addressed by several authors who have contributed to the conceptualization of the process [56, 87], but not in an holistic way. As stated by Jaeschke et al. [34], there are several works in information visualization that overlap, but none of them integrates all these aspects. To fill this gap, they propose the Information Visualization Modelling Language (IVML), where the vocabulary is made of the concepts related to pipeline, data features, data types, cognition, interaction, tasks, visualization techniques, multiple views, visual representations, and visual features. In addition, IVML describes the relationships between these concepts through the grammar. Jaeschke et al. also emphasize the importance of creating a language that meets some requirements, like verifiability, unambiguity, expressiveness, extensibility, or scalability. Our proposal, on the contrary, provides designers and developers with a visual model allowing to fully describe an information visualization system through concepts with different levels of abstraction and a series of generalization, association, and composition relationships. Nonetheless, our model has also been created to be unambiguous, expressive, verifiable, and especially extensible and scalable.

8 Conclusion and future work

We have presented a conceptual model of information visualization that graphically illustrates, in a standard notation (UML), which are the main concepts that have to be considered when designing an information visualization system, and how they relate.

Even if we have only provided examples in a given domain (research-oriented information-seeking), we believe the model is easily adaptable to any domain, as we have provided it with high levels of abstraction, scalability and flexibility.

As a result, the presented model can be used as a framework to exhaustively describe or compare any existing information visualization systems, but also to guide the design of a new information visualization system as it can be used as structured and semantically rich checklist of the concepts and relationships that have to be taken into account.

As a continuation of this work, we plan apply the model in different domains. In addition, we plan to assess the usefulness of the model for the comparison of information visualization systems by performing a thorough comparison of the most relevant current ones. Finally, the model will be used as a guideline to design a new information visualization system supporting the information-seeking process performed by researchers. In all cases, the involved designers, developers, and experts participating in the description, comparison, or creation of the information visualization systems will be asked to evaluate the model quantitatively and qualitatively. The evaluation will allow collecting both objective and subjective measures, like the time required to create a specific model, the level of coverage, and the existence of concepts and/or relationships in a system that cannot be represented in the model (objective measures), and the users' opinion, perception of use, generated feelings, and degree to which they would like to use the model in the future (subjective measures).

Acknowledgments We would like to thank the anonymous reviewers for critically reading this manuscript. Their comments and suggestions have helped to improve the paper to a great extent.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Aaltonen A, Lehikoinen J (2005) Refining visualization reference model for context information. Pers Ubiquitous Comput 9(6):381–394. https://doi.org/10.1007/s00779-005-0349-4
2. Agus M, Bettio F, Giachetti A, Gobbetti E, Iglesias Guitián JA, Marton F, Nilsson J, Pintore G (2009) An interactive 3D medical visualization system based on a light field display. Vis Comput 25(9):883–893. https://doi.org/10.1007/s00371-009-0311-y
3. Ahlberg C, Shneiderman B (1994) Visual information seeking: tight coupling of dynamic query filters with Starfield displays. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems- CHI '94*, Boston, pp 313–317. https://doi.org/10.1145/191666.191775
4. Amar RA, Stasko JT (2004) A knowledge task-based framework for design and evaluation of information visualizations. In IEEE Symposium on Information Visualization, 2004 (INFOVIS 2004), pp 143–150. https://doi.org/10.1109/INFVIS.2004.10
5. Amar R, Eagan J, Stasko J (2005) Low-level components of analytic activity in information visualization. In IEEE Symposium on Information Visualization, 2005 (INFOVIS 2005), pp 111–117. https://doi.org/10.1109/INFVIS.2005.1532136

6. Anderson E (1936) The species problem in Iris. Ann Mo Bot Gard 23(3):457–509. https://doi.org/10.2307/2394164
7. Behrisch M, Blumenschein M, Kim NW, Shao L, el-Assady M, Fuchs J, Seebacher D, Diehl A, Brandes U, Pfister H, Schreck T, Weiskopf D, Keim DA (2018) Quality metrics for information visualization. Computer Graphics Forum 37(3):625–662. https://doi.org/10.1111/cgf.13446
8. Benford S, Mariani J (1995) Virtual Environments for Data Sharing and Visualisation — Populated Information Terrains. In: Interfaces to Database Systems (IDS94). Workshops in Computing, pp 168–182. https://doi.org/10.1007/978-1-4471-3818-1_9
9. Benford S, Snowdon D, Greenhalgh C, Ingram R, Knox I, Brown C (1995) VR-VIBE: a virtual environment for co-operative information retrieval. Computer Graphics Forum 14(3):349–360. https://doi.org/10.1111/j.1467-8659.1995.cgf143_0349.x
10. Bertin J (1981) Graphics and graphic information processing. Walter de Gruyter
11. Blach R, Wenzel G, Dangelmaier M, Frohnmayer J (2007) Towards a Physical Based Interaction-Model for Information Visualization. In: *Universal Access in Human-Computer Interaction. Ambient Interaction*. Springer, Berlin, pp 750–758
12. Bostock M, Heer J (2020) Data-driven documents (D3). https://d3js.org. Accessed 26 August 2020
13. Bouali F, Guettala A, Venturini G (2016) VizAssist: an interactive user assistant for visual data mining. Vis Comput 32(11):1447–1463. https://doi.org/10.1007/s00371-015-1132-9
14. Brehmer M, Munzner T (2013) A multi-level typology of abstract visualization tasks. IEEE Trans Vis Comput Graph 19(12):2376–2385. https://doi.org/10.1109/TVCG.2013.124
15. Bugajska M (2003) Spatial visualization of abstract information', PhD Thesis, Technische Wissenschaften ETH Zürich. http://e-collection.library.ethz.ch/view/eth:26330 (accessed Aug. 26, 2020)
16. Bugajska M (2005) Framework for spatial visual design of abstract information. Proc Int Confer Information Visualis 2005:713–723. https://doi.org/10.1109/IV.2005.51
17. Card SK, Mackinlay J (1997) The structure of the information visualization design space. In: Proceedings of VIZ '97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium, pp 92–99. https://doi.org/10.1109/INFVIS.1997.636792
18. Card SK, Robertson GG, Mackinlay JD (1991) The information visualizer, an information workspace. Proceedings of the SIGCHI conference on Human factors in computing systems Reaching through technology - CHI '91, New York, pp 181–186. https://doi.org/10.1145/108844.108874
19. Card SK, Mackinlay JD, Shneiderman B (1999) Readings in information visualization: using vision to think. Morgan Kaufmann Publishers Inc., San Francisco
20. Carr DA (1999) Guidelines for designing information visualization. In: Proceedings of the Ericsson Conference on Usability Engineering 1999 (ECUE'99), pp 1–7
21. Chi EH (2000) A taxonomy of visualization techniques using the data state reference model', in *IEEE Symposium on Information Visualization* 2000. INFOVIS 2000. Proceedings, pp 69–75. https://doi.org/10.1109/INFVIS.2000.885092.
22. Craig P, Huang X, Chen H, Wang X, Zhang S (2015) Pervasive information visualization: toward an information visualization design methodology for multi-device co-located synchronous collaboration. In: 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing. Pervasive Intelligence and Computing, Liverpool, pp 2232–2239. https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.330
23. Davies J (2014) Countries and Antarctica projected using the conformal stereographic projection, rescaled and sorted by land area. https://www.jasondavies.com/maps/countries-by-area/. Accessed 26 Aug 2020
24. De Antonio A, Moral C, Klepel D, Abente MJ (2012) Gesture-based interaction with 3D visualizations of document collections for exploration and search
25. De Antonio A, Moral C, Klepel D, Abente MJ (2013) 3D gesture-based exploration and search in document collections. In: Proceedings of the 17th international conference on electronic publishing, pp 13–22. https://doi.org/10.3233/978-1-61499-270-7-13
26. de Antonio A, Moral C, Klepel D, Abente MJ (2013) Gesture-based control of the 3D visual representation of document collections for exploration and search. Inf Serv Use 33(2):139–159
27. Einsfeld K, Agne S, Deller M, Ebert A, Klein B, Reuschling C (2006) Dynamic visualization and navigation of semantic virtual environments. In: Proceedings of the International Conference on Information Visualisation 2006 (IV'06), pp 569–574. https://doi.org/10.1109/IV.2006.46
28. Fairchild KM, Poltrock SE, Furnas GW (1988) SemNet: three-dimensional graphic representation of large knowledge bases. In: Guindon R (ed) Cognitive science and its applications for human–computer interaction. Lawrence Erlbaum Associates, Hillsdale, pp 201–233
29. Fernandez R, Fetais N (2018) Information Visualization Techniques for Building Better Visualization Models. *Recent Trends in Computer Applications*, Cham, pp 279–296. https://doi.org/10.1007/978-3-319-89914-5_17

30. Google Inc (2020) About Google scholar. https://scholar.google.com/intl/es/scholar/about.html. Accessed 26 Aug 2020
31. Healey CG, Booth KS, Enns JT (1996) High-speed visual estimation using preattentive processing. ACM Trans Comput Hum Interact 3(2):107–135. https://doi.org/10.1145/230562.230563
32. Heer J, Card SK, Landay JA (2005) prefuse: a toolkit for interactive information visualization, *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '05*, New York, p 421. https://doi.org/10.1145/1054972.1055031
33. Irshad S, Rambli DRA, Sulaiman SB (2019) An interaction design model for information visualization in immersive augmented reality platform. Proceedings of the 17th International Conference on Advances in Mobile Computing & Multimedia, New York, pp 200–206. https://doi.org/10.1145/3365921.3365939
34. Jaeschke G, Leissler M, Hemmje M (2005) Modeling interactive, 3-dimensional information visualizations supporting information seeking behaviors. Lect Notes Comput Sci 3426:119–135
35. Järvelin K, Wilson TD (2003) On conceptual models for information seeking and retrieval research. Information Res 9(1):163
36. Joblove GH, Greenberg D (1978) Color spaces for computer graphics. ACM SIGGRAPH Comput Graphics 12(3):20–25. https://doi.org/10.1145/965139.807362
37. Kirste T, Rauschenbach U (1996) A presentation model for mobile information visualization. Comput Graph 20(5):669–681. https://doi.org/10.1016/S0097-8493(96)00041-6
38. Liu Z, Stasko JT (2010) Mental models, visual reasoning and interaction in information visualization: a top-down perspective. IEEE Trans Vis Comput Graph 16(6):999–1008. https://doi.org/10.1109/TVCG.2010.177
39. Liu S, Cui W, Wu Y, Liu M (2014) A survey on information visualization: recent advances and challenges. Vis Comput 30(12):1373–1393. https://doi.org/10.1007/s00371-013-0892-3
40. Lu M, Chen S, Lai C, Lin L, Yuan X (Nov. 2017) Frontier of information visualization and visual analytics in 2016. J Vis 20(4):667–686. https://doi.org/10.1007/s12650-017-0431-9
41. Luo Y, Qin X, Tang N, Li G (2018) DeepEye: Towards Automatic Data Visualization. In: 2018 IEEE 34th international conference on data engineering (ICDE), pp 101–112. https://doi.org/10.1109/ICDE.2018.00019
42. MacEachren AM (1995) How Maps Work: Representation, Visualization, and Design, vol 3. The Guilford Press
43. Mackinlay JD (1986) Automating the design of graphical presentations of relational information. ACM Trans Graph 5(2):110–141. https://doi.org/10.1145/22949.22950
44. Mackinlay JD, Robertson GG, Card SK (1991) The perspective wall. Proceedings of the SIGCHI conference on Human factors in computing systems Reaching through technology - CHI '91, New York, pp 173–176. https://doi.org/10.1145/108844.108870
45. Mazza R (2009) Introduction to information visualization. Springer-Verlag
46. McNabb L, Laramee RS (2017) Survey of surveys (SoS) - mapping the landscape of survey papers in information visualization. Comput Graphics Forum 36(3):589–617. https://doi.org/10.1111/cgf.13212
47. Mei H, Ma Y, Wei Y, Chen W (2018) The design space of construction tools for information visualization: a survey. J Vis Lang Comput 44:120–132. https://doi.org/10.1016/j.jvlc.2017.10.001
48. Mei H, Chen W, Ma Y, Guan H, Hu W (2018) VisComposer: a visual programmable composition environment for information visualization. Visual Informatics 2(1):71–81. https://doi.org/10.1016/j.visinf.2018.04.008
49. Meyer M, Sedlmair M, Quinan PS, Munzner T (2015) The nested blocks and guidelines model. Inf Vis 14(3):234–249. https://doi.org/10.1177/1473871613510429
50. Momma K, Izumi F (2008) VESTA: a three-dimensional visualization system for electronic and structural analysis. J Appl Crystallogr 41(3):653–658. https://doi.org/10.1107/S0021889808012016
51. Moral C (2016) Modeling the visualization and exploration of document collections with user and purpose-based adaptation. Ph.D Dissertation, Universidad Politécnica de Madrid. http://oa.upm.es/44561/. Accessed 26 Aug 2020
52. Moral C, De Antonio A, Ferre X, Lara G (2015) A coding system for qualitative studies of the information-seeking process in computer science research. Inf Res 20(4):699. http://www.informationr.net/ir/20-4/paper699.html#.VpfgN5PhC-p. Accessed 26 Aug 2020
53. Moral C, De Antonio A, Ferre X (2017) A visual UML-based conceptual model of information-seeking by computer science researchers. Inf Process Manag 53(4):963–988. https://doi.org/10.1016/J.IPM.2016.10.005
54. Morgan R, Grossmann G, Schrefl M, Stumptner M, Payne T (2018) VizDSL: A Visual DSL for Interactive Information Visualization. *Advanced Information Systems Engineering*, Cham, pp 440–455. https://doi.org/10.1007/978-3-319-91563-0_27

55. Munzner T (2009) A nested process model for visualization design and validation. IEEE Trans Vis Comput Graph 15(6):921–928. https://doi.org/10.1109/TVCG.2009.111
56. Munzner T (2014) Visualization analysis and design. A K Peters/CRC Press
57. Nowell L, Schulman R, Hix D (2002) Graphical encoding for information visualization: an empirical study. In: IEEE Symposium on Information Visualization, 2002. INFOVIS 2002, pp 43–50. https://doi.org/10.1109/INFVIS.2002.1173146
58. Object Management Group (2017) About the unified modeling language specification version 2.5.1. Unified Modeling Language. https://www.omg.org/spec/UML. Accessed 26 Aug 2020
59. Oliveira EC, Oliveira LC, Cardoso A, Mattioli L, Lamounier EA (2017) Meta-model of information visualization based on treemap. Universal Access in the Information Society 16(4):903–912. https://doi.org/10.1007/s10209-016-0477-9
60. Perez C, de Antonio A (2003) An experimental evaluation of Spatialization algorithms for Visualising text corpora. Workshop on Knowledge and Information Visualization (KIV'03), 3rd International Conference on Knowledge Management (I-Know'03), Graz, pp 103–108
61. Perez C, de Antonio A (2004) 3D visualization of text collections: an experimental study to assess the usefulness of 3D. Proc Eighth Int Confer Inform Visualis IV:317–323. https://doi.org/10.1109/IV.2004.1320163
62. Pike WA, Stasko J, Chang R, O'Connell TA (2009) The science of interaction. Inf Vis 8(4):263–274. https://doi.org/10.1057/ivs.2009.22
63. Preece J, Rogers Y, Sharp H (2015) *Interaction Design: Beyond Human-Computer Interaction*, 4th ed. Wiley
64. Rekimoto J, Green M (1993) The Information Cube: Using Transparency in 3D Information Visualization. In: Proceedings of the Third Annual Workshop on Information Technologies & Systems (WITS'93), pp 125–132
65. Ren L, Tian F, Zhang X, Zhang L (2010) DaisyViz: a model-based user interface toolkit for interactive information visualization systems. J Vis Lang Comput 21(4):209–229. https://doi.org/10.1016/j.jvlc.2010.05.003
66. Richter H, März L (2000) Toward a standard process: the use of UML for designing simulation models. In Proceedings of the 32nd Winter Simulation Conference (WSC'00), pp 394–398. https://doi.org/10.1109/WSC.2000.899744
67. Risch JS, Rex DB, Dowson ST, Walters TB, May RA, Moon BD (1997) The STARLIGHT information visualization system. In: Information Visualization, 1997. Proceedings., 1997 IEEE Conference on, pp 42–49. https://doi.org/10.1109/IV.1997.626486
68. Robertson GG, Mackinlay JD, Card SK (1991) Cone trees: animated 3D visualizations of hierarchical information. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New Orleans, pp 189–194. https://doi.org/10.1145/108844.108883
69. Roth RE (2013) An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. IEEE Trans Vis Comput Graph 19(12):2356–2365. https://doi.org/10.1109/TVCG.2013.130
70. Satyanarayan A, Wongsuphasawat K, Heer J (2014) 'Declarative interaction Design for Data Visualization. *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, New York, pp 669–678. https://doi.org/10.1145/2642918.2647360
71. Satyanarayan A, Russell R, Hoffswell J, Heer J (2016) Reactive Vega: a streaming dataflow architecture for declarative interactive visualization. IEEE Trans Vis Comput Graph 22(1):659–668. https://doi.org/10.1109/TVCG.2015.2467091
72. Schulz HJ, Nocke T, Heitzler M, Schumann H (2013) A design space of visualization tasks. IEEE Trans Vis Comput Graph 19(12):2366–2375. https://doi.org/10.1109/TVCG.2013.120
73. Semiu AA, Jamaludin Z (2013) Information visualization techniques usage model', Proceedings of the International Conference on Computing & Informatics, pp 247–252
74. Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings 1996 IEEE Symposium on Visual Languages, pp 336–343. https://doi.org/10.1109/VL.1996.545307.
75. Shneiderman B, Feldman D, Rose A, Ferre X (2000) Visualizing digital library search results with categorical and hierarchical axes. Proceedings of the 5th ACM Conference on Digital libraries - DL '00, pp 57–66. https://doi.org/10.1145/336597.336637.
76. Shneiderman B, Plaisant C, Cohen M, Jacobs S (2009) Designing the user Interface: strategies for effective human-computer interaction (5th ed.). Addison-Wesley Publishing Company, Boston
77. Siirtola H, Räihä K-J, Säily T, Nevalainen T (2010) Information visualization for corpus linguistics: towards interactive tools. *Proceedings of the first international workshop on Intelligent visual interfaces for text analysis - IVITA '10*, New York, p 33. https://doi.org/10.1145/2002353.2002365

78. Spoerri A (1993) InfoCrystal: a visual tool for information retrieval and management. In: Proceedings of the Second International Conference on Information and Knowledge Management 1993 (CIKM'93), New York, pp 11–20. https://doi.org/10.1145/170088.170095
79. Sprenger TC, Gross MH, Eggenberger A, Kaufmann M (1997) A Framework for Physically-Based Information Visualization. *Visualization in Scientific Computing '97*, Vienna, pp 71–83
80. Stasko J (2014) Value-Driven Evaluation of Visualizations, BELIV'2014, Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization, pp 46–53. https://doi.org/10.1145/2669557.2669579.
81. Stevens P, Pooley R (2006) Using UML: software engineering with objects and components, 2nd edition. Pearson Education, Essex
82. Stolte C, Tang D, Hanrahan P (2002) Polaris: a system for query, analysis, and visualization of multidimensional relational databases. IEEE Trans Vis Comput Graph 8(1):52–65. https://doi.org/10.1109/2945.981851
83. Sugibuchi T, Spyratos N, Siminenko E (2009) A framework to analyze information visualization based on the functional data model. In: Proceedings of the 13th international Conference on Information Visualisation 2009 (IV'09), pp 18–24. https://doi.org/10.1109/IV.2009.56
84. van Wijk JJ (2005) The value of visualization. In: IEEE Conference on Visualization 2005 (VIS'05), pp 79–86. https://doi.org/10.1109/VISUAL.2005.1532781
85. Wang X-M, Zhang T-Y, Ma Y-X, Xia J, Chen W (2016) A survey of visual analytic pipelines. J Comput Sci Technol 31(4):787–804. https://doi.org/10.1007/s11390-016-1663-1
86. Ware C (2020) *Information Visualization: Perception for Design*, fourth edition. Morgan Kaufmann, Cambridge
87. Wilkinson L (2005) The grammar of graphics, Second Edition. Springer Science & Business Media
88. Wiss U, Carr D, Jonsson H (1998) Evaluating three-dimensional information visualization designs: A case study of three designs', in Proceedings of the IEEE Conference on Information Visualization, pp 137–144. https://doi.org/10.1109/IV.1998.694211.
89. Wu J (2010) WIVF: Web information visualization framework based on information architecture 2.0. In: Computer and Automation Engineering (ICCAE), 2010 The 2nd international conference on, vol 1, pp 734–738. https://doi.org/10.1109/ICCAE.2010.5451266
90. Yi JS, Kang Y, Stasko JT, Jacko JA (2007) Toward a deeper understanding of the role of interaction in information visualization. IEEE Trans Vis Comput Graph 13(6):1224–1231
91. Zhang J (2001) TOFIR: a tool of facilitating information retrieval – introduce a visual retrieval model. Inf Process Manag 37(4):639–657. https://doi.org/10.1016/S0306-4573(00)00042-X
92. Zhang J, Nguyen T (2005) WebStar: a visualization model for hyperlink structures. Inf Process Manag 41(4):1003–1018. https://doi.org/10.1016/j.ipm.2004.03.005
93. Zhou MX, Chen M, Feng Y (2002) Building a visual database for example-based graphics generation. In: IEEE Symposium on Information Visualization. INFOVIS 2002, pp 23–30. https://doi.org/10.1109/INFVIS.2002.1173143

Cristian Moral received his MSc and PhD degrees in Computer Science from Universidad Politécnica de Madrid in 2011 and 2016, respectively. He is Assistant Professor at the Computer Science School at UPM from 2014. His main research interests Human-Computer Interaction in ICT for healthy ageing, Information Retrieval and Visualization, Virtual and Augmented Reality, and eHealth.

Angélica de Antonio received her Ph.D. in Computer Science in 1994 from the Universidad Politécnica de Madrid, Spain. She is Director of the laboratory "Madrid HCI Lab - Decoroso Crespo" since 1995. Her areas of interest are educational software, particularly intelligent tutoring systems; virtual and augmented reality, mainly on its application for education and training; intelligent virtual agents, specially working as pedagogical agents; human-computer interaction techniques for 3D environments; and the software engineering methods and techniques that make it possible to develop such complex systems. She is a senior lecturer of University in the Universidad Politécnica de Madrid.

Xavier Ferre received MSc and PhD degrees from Universidad Politécnica de Madrid (UPM) in 1996 and 2005, respectively. He is Associate Professor at the Computer Science School at UPM from 2008. His primary research interests are usability and User eXperience (UX) in mobile applications for older users, integration of usability practices into software engineering, and ICT applied to health care.

Jaime Ramirez received the M.Sc. degree in computer science and the Ph.D. degree in computer science from Politécnica de Madrid University (UPM), in 1996 and 2002, respectively. He is currently Associate Professor at the Computer Science School at UPM. His main research interests include ontology engineering, 3D virtual environments for education/training and adaptive systems.