2010 14th International Conference Information Visualisation

From Data to Knowledge

Visualizations as transformation processes within the Data-Information-Knowledge continuum

Luca Masud, Francesca Valsecchi, Paolo Ciuccarelli, Donato Ricci, Giorgio Caviglia
INDACO Department, Density Design Research Group
Politecnico di Milano
Milan, Italy
lucamasud@gmail.com, francesca.valsecchi@polimi.it, paolo.ciuccarelli@polimi.it, donato.ricci@gmail.com, giorgio.caviglia@gmail.com

***Abstract*—In this paper we suggest a different approach that considers visualizations in the wider domain of communication and defines a model capable of taking into account the context in which visualizations act as communication tools. In this perspective we consider visualizations as transformation processes within the Data-Information-Knowledge (DIK) continuum. In the paper we discuss the continuum, and apply the transformation process model to the main disciplines of visualization. Visualizations represent powerful cognitive tools that surround our everyday life [1]. By doing this we are able to think about visualization from a multidisciplinary perspective, exploring the role of visualizations in design practice, as artifacts that are used in facing problems of various degrees of complexity and nature. The contribution of the model is mainly addressed to orientate and nurture the reflective practice and to formalize the strategic more than the technical role of visualizations in the design discipline.**

Information Visualization, DIK, Data, Information, Knowledge, Knowledge continuum, Design, transformations, processes

I. THE DESIGN SENSITIVITY TO VISUALIZATIONS

In our everyday life we are surrounded by many kind of visualizations: from popular Infographics and customized maps available on newspapers, magazines or websites to the more specific and complex ones inside working environment for the explanation of workflows, complex processes or organizational structures.

This ever-growing diffusion of visualization tools even in non-expert contexts and during decision making processes and planning phases (opposed to an exclusively analytical approach) requires a more open way of thinking about visualizations, able to take in account the "world out there" [2], the "'real' world" [3]. Different disciplines (especially Information Visualization) have deeply investigated visualizations mainly from a strictly analytical point of view, focusing their reflections on some particular kind of visualization, used especially in very narrow, peculiar and usually very technical contexts.

While Tuftean authors advocate the need of a strictly functionalist and analytic approach to visualization in the domain of languages and techniques, other authors [3] [4] [5] wrote about a more communicative approach to visualization (casual info-vis, "information visualization for the people", information-aesthetics). Even though both these positions are useful to reflect about specific kinds of visualizations, they seem to be not enough (o not completely effective) to support a interdisciplinary design-oriented approach. As a matter of fact, there is not still a common ground model to rely upon.

In a perspective in which we have the opportunity of using visualization to face problems of various degrees of complexity and different nature, communication design can play in fact a useful role [6]. This may be especially true if "the challenges of the modern world require integrative problem solving and, at a more comprehensive level, holistic thought and transdisciplinary schema promote unity of knowledge" [7] and given that "without integrative disciplines of understanding, communication, and action, there is little hope of sensibly extending knowledge beyond the library or laboratory in order to serve the purpose of enriching human life" [8]. It is exactly in this way of thinking that design – "by nature an interdisciplinary, integrative discipline" [9] – and "the ability of designers to discover new relationships among signs, things, actions, and thoughts [...]" [8] can play a big role.

Working on visualizations from a communication design perspective requires a different approach, one able to consider them in the wider domain of communication strategy. This is the reason why we need a model able to cope with the many nuances of visualizations, a model capable of taking into account the context in which visualizations act as communication tools.

In such a design perspective we must start to refer to visualizations as means to achieve purposes. This does not refer strictly to the idea of representing of high-dimensional sets of data: "the ability to visualize complex information does not refer solely to the communication of quantitative information but it also deals with the visual narration of values and qualitative data."[10]. So the aims of a

1550-6037/10 $26.00 © 2010 IEEE
DOI 10.1109/IV.2010.68

visualization can include indeed the capability to make sense of context, communicate impressions, telling stories. We need to change our perspective so that visualizations are not merely defined by the technology they involve (as the Information Visualization definition declares – see Paragraph III), but rather by the relation with the aim and context they are designed for and the recipient they want to reach.

In a user-centered design approach visualizations can be seen as *problem-solving tools* differently deployed according to their efficiency in a given context and for a specific target. This leads to different kinds of languages and approaches in the visualization domain.

The actual classification of visualizations based on the different disciplinary contexts they arise from (see Paragraph III) might be good to describe a visualization within a single context but it fails in describing them in a cross-boundaries perspective.

By the design framework visualizations are tools just in the sense that they include users, context and aims in their structure and they are expected to orientate further knowledge and practice. Design is not pretending to provide univocal position; as designers we need to start from the purposes of communication focusing the reflection on the processes rather than the outputs. Therefore we suggest that visualizations can be categorized by the communicational aspects they imply: what, in which way and to whom communicate. From a designer's perspective visualizations represent the process that moves from data to knowledge, where each visualization is seen as a transformation artifact within the data-information-knowledge continuum (DIK) [11]. In the next paragraphs we are going to develop this concept.

II. Description of DIK Continuum

Since visualizations work with data, information and knowledge (as in some way it's hinted – though in a misleading way, since there's no direct correlation between the discipline's name and the visualization per se – by the main discipline's name: Data Visualization, Information Visualization, Knowledge Visualization) we propose a framework based on the Data-Information-Knowledge continuum. As Bellinger, Castro and Mills [11] refer, the continuum starts with raw data, "it simply exists and has no significance beyond its existence (in and of itself). It can exist in any form, usable or not", proceeds with information that occurs when data has been given meaning by way of relational connections. Also in this case, "This "meaning" can be useful, but does not have to be". Finally knowledge is the appropriate collection of information, "such that its intent is to be useful" [11]. For example data may be the different temperatures measured by a thermometer. By relating these different temperatures it is possible to understand if the temperature is rising or decreasing, giving the data a meaning and thus transforming it in information (such that it may be useful). Lastly, the fact that it is known that the temperature is decreasing may trigger an action: e.g. if the temperature was measured at home then the heating might be turned on.

In this perspective visualizations are not merely final outcomes of representing data, information and knowledge. Instead they have to be conceived as transformation processes within the DIK continuum (figure 1). Visualizations are in fact able to gather data, information or knowledge (*materials*), visualizing them in an *artifact* is information, to eventually create new knowledge (*objective*) in the recipient. In Paragraph IV we will see how different kinds of visualization create different kinds of knowledge. The process consists of two main parts: producer's and user's. The first one is the act of designing a visualization: just as information is selecting, ordering and relating data a visualization is always selecting and ordering. Visualizing means deciding what and how to show of a given data set or information. For this reason visualizations are always information in the "universal domain"[12]. The second part of the process is related to the user's interaction with the visualization. As such the results of this part are not totally controllable by the producer.

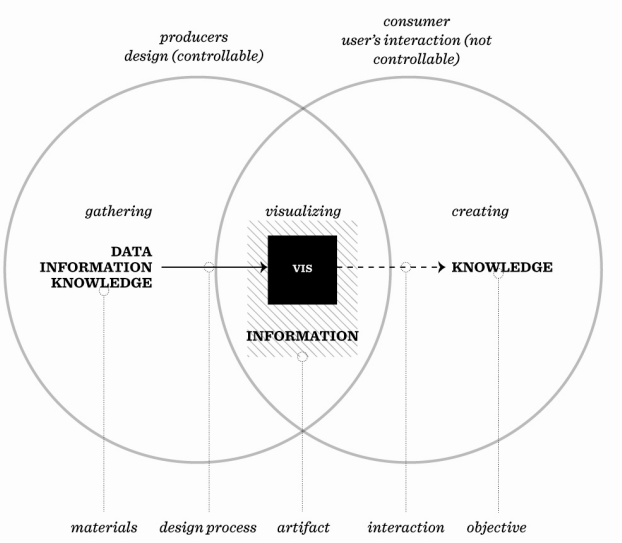

Figure 1. Visualizations as processes within the DIK continuum

III. Visualization Disciplines and Approaches

Before understanding the different roles of visualization it is better to start identifying the most important disciplines in which visualization takes form. The following list is the result of a research developed to make the various disciplines' domains clear. As a matter of fact there are currently no defined boundaries between the different disciplines and their names are sometimes used interchangeably. Here are presented the main and most important disciplines (or domains) of visualization: even if we are aware that this list can be improved or completed it serves its purpose of illustrating the current state of visualization's different approaches. Visualizations in the art domain are not discussed in this paper. In fact art may create

awareness around a given topic but it is not always able and it is not created to give insights or explain a phenomenon.

a) Data Visualization: it refers to the practice of using graphical representation – "information which has been abstracted in some schematic form" [13] – to provide visual insights in sets of data. It may refer both to static and dynamic representations (thus covering Information Visualization too).

b) Information Visualization: "The use of computer-supported, interactive, visual representations of abstract data to amplify cognition" [14]. While Data Visualization is usually used as an umbrella term describing every form of visual representation of data, Information Visualization is restricted to computer-supported visualizations.

c) Scientific Visualization: "Scientific visualization [...] deals with physically-based data. This kind of data is defined in reference to space coordinates, which makes it relatively easy to visualize in an intuitive way. The space coordinates in the dataset are mapped to screen coordinates. Examples are geographic data and computer tomography data of a body." [15]

d) Information Aesthetics: "Information aesthetics forms a cross-disciplinary link between information visualization and visualization art. It adopts more interpretive mapping techniques to augment information visualization with extrinsic meaning, or considers functional aspects in visualization art to more effectively convey meanings underlying datasets.[9]

e) Infographics: although "Information Graphics" refers to the tools and techniques involved in the graphical representation of data [16] in this paper the term is used in relation to newspaper's infographics; for this reason is important to remember that this kind of infographics' background, unlike data visualization, stands in journalism, design and the art of telling stories. Newspaper infographics story is almost as old as newspapers themselves [17] and are a great example of visually communicating informations to a broad, non-expert target.

f) Knowledge Visualization: Knowledge Visualization unlike Information Visualization uses visual representation to trasfer knowledge between at least to person or a group or persons rather than for data insights [18]. Given the kind of applications of Knowledge Visualization artefacts (Knowledge Management. Organizational Science) the experiential and actuative – getting someone to get action – dimension is the main feature of this discipline.

IV. Visualizations Acting in the DIK Continuum

We can note from the previous descriptions that the boundaries between disciplines are quite blurred and that a new approach based on aim and targets can help the reasoning about languages and medias involved.

In this paragraph we will integrate the Data-Information-Knowledge continuum with the different fuctions a visualization can carry out.

Starting from the list proposed on the previous paragraph we are able to delineate three main blocks of visualizations, each of them based on the kind of materials and aims the visualization works with. Furthermore we can distinguish different knowledge kinds resulting from the visualization process.

It could be useful to briefly outline the knowledge types discussed in this paper. We will in fact distinguish three kind of knowledge: declarative, procedural and conditional [19][20]. Even if authors like Nonaka [21] distinguish five types of knowledge (declarative, procedural, causal, locational, relational) for our purposes the tripartition suffices: a more refined analysis might work well in distinguish different kinds of visualization artefacts – e.g. locational knowledge for georeferenced visualizations – but do not work so well for the macro-categories of this paper.

The first one includes visualizations that take data (abstract or not) and convert it into information, in order to let the user know something and make assumptions on the data. It's a transformation that explores the *know-what* (or *know about*), to which we refer as declarative knowledge.

The second one refers to those visualizations, (i.e. newspaper Infographics and Information Aesthetics Visualizations) that do not only gather data and visualize it but are also able to "tell a story" and communicate pieces of crystallized information (as opposed to raw data). This information can then be used by the recipient to understand something or to know how to do something (e.g. instruction manual pictograms, Infographics about procedures). For this reason the second group does not only work on declarative knowledge but on procedural one too. We should note that the diagram in figure 2 shows data/knowledge instead of data/information because understanding a theme is fundamental for being able to talk about it. The internalization of information becomes knowledge [22].

Lastly the third group does not take data as the starting point. In fact, the aim of Knowledge Visualization is transferring knowledge in a collaborative context, not making sense in a high-dimensional data set. Knowledge Visualizations does not only communicate *how* to do things (procedural knowledge) but they are also able to transfer knowledge of *when* and *why* the recipient should use his knowledge (conditional knowledge [19]). In this perspective the ability to "take action" in a context is fundamental in knowledge transfer.

We will call these three categories **analytical** visualizations, **communicative** visualizations and **formative** visualizations.

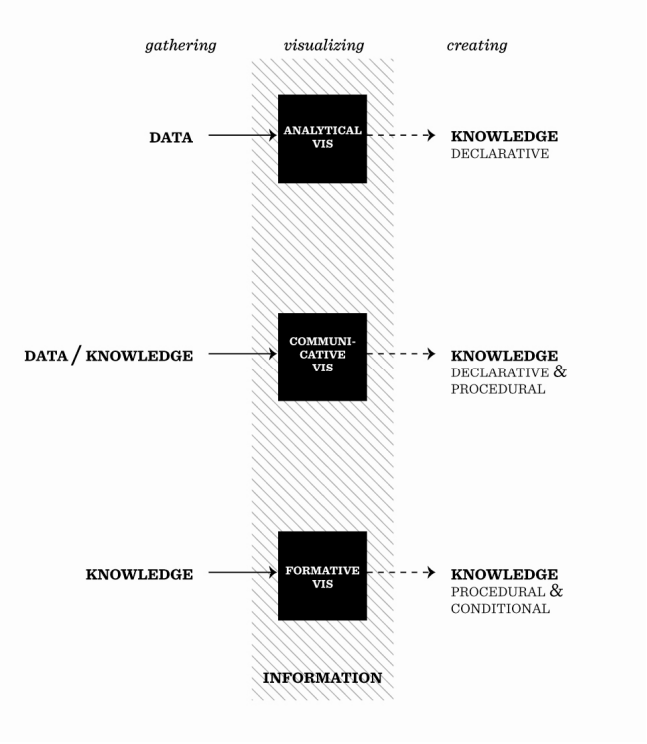

Figure 2. Visualizations as processes within the DIK continuum

This three kinds of visualization have different recipients and contexts of use. Therefore, their function, aim and language greatly differ from one another.

Although a deeper analysis is needed here, we start from the identification of some main guidelines for each type of visualization.

A. *Analytical Visualizations*

In their functionalist approach analytical visualizations largely use reversible techniques so that the recipient is able to infer data from the representation (see Lau and Vande Moere [9]). They are usually deployed for a strictly technical target who use those visualization for work: this kind of user have to rely on visualizations in their everyday work. By technical here we do not intend it in a visualization perspective but simply professionals who utilizes visual representations: e.g. analysts, economists, statistical scientists, medical personnel, etc.

B. *Communicative Visualizations*

In this case visualizations are not used for a detailed data analysis but for storytelling or to communicate the meaning of the relations of data: the author acts as intermediary in the communication process so that the visual representation is not used to make assumptions and analysis but to communicate results. As Pousman, Stasko and Mateas notes [8] this kind of visualizations (which they call "casual information visualization") are able to provide awareness, social and reflective insights, even though they do not offer analytical ones. This approach is well explained by the different kind of recipient of communicative visualizations: being this kind of representations deployed on mass media the general audience to which they are addressed may not have the proficiencies for such kind of visualization. In this group a pictogrammatic language and the use of metaphors and illustrations are usually employed in the visualization, since an accurate data retrieval is not its main aim: chances are that a data-ink ratio approach to visualization may prove non-effective or even counter-productive in this context.

C. *Formative Visualizations*

This kind of visualizations are deeply characterized by the context in which they are used. They share in fact techniques with communicative visualizations even though they have a different target and context of use.

They are used as a support in knowledge transfer inside cooperative work groups. They represent workflows, processes and they are able to instruct users of their role in that cooperative context. The real distinctive feature of formative visualizations is action: those visualizations are realized for people in active roles inside an organization or work group so that they are able to know how, when and why to act in a given context. Knowledge may in fact be described as "actionable information" [23] and some kind of knowledge are strictly bound to experience (i.e. Procedural knowledge that is "knowledge about how, when and why to do something" [24])

V. DISCUSSION & CONCLUSIONS

As we highlighted visualizations may be considered as a process. Considering aims, targets and contexts we defined three different processes which, while sharing the same base structure, differ in the final outcome. In fact the artifacts belonging to each of these three processes use different kind of languages and techniques according tohe different transformation processes within the continuum.

This kind of classification is mainly addressed to orientate and nurture the reflective practice and to formalize the strategic more than technical role of visualizations in the design discipline.

Even if the interest for data, information and knowledge visualization grows at an impressive pace, the possibility to define a wider and strategic role for these competences in a knowledge society - especially within decision-making processes related to complex problems - suffers from a fragmented disciplinary field.

As the problems we are facing became more and more complex and require a interdisciplinary approach and a new framework for knowledge processes, design can play a specific and important role. Exploiting their form-giving ability and the attitude to connect disciplines, designers are able to produce - together with other experts - visualizations that respond to the multiple targets and goals of the different phases that articulate complex-problem solving or decision making processes. To accomplish this role, a framework is needed to connect the different visualization disciplines, and to overcome the limits of the traditional functional or typological taxonomies.

We identified the Data, Information, Knowledge continuum as the key element for a framework that considers visualization as a process and not as a product. As a matter of fact we are involved in the practices (the process) leading to the visualization artifact (the product) more than in the artifact in itself. Considering the three different approaches to visualization, the passage from product to process allows to define the visualization techniques and languages according to the proposed objectives, targets and contexts in which they will be deployed.

In our experience, it shows its usefulness in defining the visualization requirements and the disciplinary partnerships needed for any of the possible context-target-goal combinations that characterize the different phases of a decision-making/problem solving process.

More research and work is needed to validate the framework in a wider number of contexts and to deeply define the specific combination of techniques and languages to be used in the three – analytical, communicative, formative – visualization domains.

REFERENCES

[1] G. Bonsiepe. "Design as tool for cognitive metabolism, the role of design in the socialisation of knowledge". In International Symposium "Ricerca+ Design". 2000.

[2] S. Few. "InfoVis as Seen by the World Out There: 2007 in Review". In IEEE Symposium on Visual Analytics Science and Technology, 2007. VAST 2007.

[3] M. Danziger. Information visualization for the people. Massachusetts Institute of Technology. 2008.

[4] Z. Pousman, J. T Stasko and M. Mateas. "Casual Information Visualization: Depictions of Data in Everyday Life". IEEE Transactions on Visualization and Computer Graphics 13, no. 6: 114. 2007.

[5] A. Lau, and A. V. Moere. "Towards a model of information aesthetics in information visualization". In Proceedings of the 11th International Conference Information Visualization, 87-92. IEEE Computer Society Washington, DC, USA. 2007.

[6] P. Ciuccarelli, D. Ricci, and F. Valsecchi. "Handling changes trough diagrams: Scale and grain in the visual representation of Complex Systems". In Cipolla C. and Peruccio P. P. Changing the change Proceedings. Turin: Allemandi, July 10. 2008.

[7] J.T. Klein. Crossing boundaries: Knowledge, disciplinarities, and interdisciplinarities. Charlottsville: The University of Virginia Press. 1996.

[8] R. Buchanan "Wicked Problems in Design Thinking". Design Issues, 8:2, 5-21, Cambridge, MA: The MIT Press. 199

[9] K. Friedman. "Theory construction in design research: criteria: approaches, and methods", Design Studies 24:6, November. 2003.

[10] G. Scagnetti, D. Ricci, G. Baule, P. Ciuccarelli. "Reshaping Communication Design Tools" Complex Systems Structural Features For Design Tools, in IASDR07: International Associations of Societies of Design Research" The Hong Kong Polytechnic University, Hong Kong, 20 (digital). 2007.

[11] G. Bellinger, D. Castro and A. Mills. "Data, information, knowledge, and wisdom". URL: http://www.systems-thinking.org/dikw/dikw.htm. 2004.

[12] C. Zins, "Conceptual approaches for defining data, information, and knowledge". Journal of the American Society for Information Science and Technology 58, no. 4: 479–493. 2007.

[13] M. Friendly, and D. J. Denis. "Milestones in the history of thematic cartography, statistical graphics, and data visualization". 2004.

[14] S. K. Card, J. D Mackinlay and B. Shneiderman. Readings in information visualization: using vision to think. Morgan Kaufmann. 1999.

[15] R. Voigt. "An extended scatterplot matrix and case studies in information visualization". Unpublished Masters thesis. Virtual Reality and Visualization Research Center, Vienna, AU. 2002.

[16] R. L. Harris. Information graphics: A comprehensive illustrated reference. Oxford University Press, USA. 1999.

[17] A. Cairo. "Sailing to the Future: Infographics in the Internet Era." Creative Commons License. University of North Carolina. 2005.

[18] R. A. Burkhard. "Learning from architects: the difference between knowledge visualization and information visualization". IV 2004. Proceedings. Eighth International Conference on, 519-524. 2004.

[19] G. Schraw. "Promoting general metacognitive awareness". Instructional Science 26, no. 1: 113–125.1998

[20] G. Judelman. Knowledge Visualization. Problems and Principles for Mapping the Knowledge Space. M. Sc. Thesis. 2004.

[21] I. Nonaka. "A dynamic theory of organizational knowledge creation". Organization science 5, no. 1: 14-37. 1994.

[22] C. Zins. "Conceptual approaches for defining data, information, and knowledge". Journal of the American Society for Information Science and Technology 58, no. 4: 479–493. 2007.

[23] E. M.Awad and H. M. Ghaziri. Knowledge Management. Upper Saddle River, NJ: Prentice Hall Publishing, Pearson Education, Inc. 2004.

[24] N. Basjes. "A model for procedural knowledge". Universiteit Nyenrode: Breukelen, The Netherlands. 2002.