PJIM PARSONS JOURNAL FOR INFORMATION MAPPING

THE PARSONS INSTITUTE FOR INFORMATION MAPPING
68 Fifth Avenue New York, NY 10011
212 229 6825 piim.newschool.edu

# *Representing the Digital Humanities Community: Unveiling The Social Network Visualization of an International Conference*

DARIO RODIGHIERO

**KEYWORDS** *Actor-network theory, data visualization, interpretant, semiotics, social network visualization*

**ABSTRACT** This paper deals with the sense of representing both a new domain as Digital Humanities and its community. Based on a case study, where a set of visualizations was used to represent the community attending the international Digital Humanities conference of 2014 in Lausanne, Switzerland, the meaning of representing a community is investigated in the light of the theories of three acknowledged authors, namely Charles Sanders Peirce for his notion of the interpretant, Ludwig Wittgenstein for his insights on the use of language, and finally Bruno Latour for his ideas of representing politics.

There results a proposal to designing and interpreting social network visualizations in a more thoughtful way, while remaining aware of the relation between objects in the real world and their visualizations. As this type of work pertains to a wider scope, we propose bringing a theoretical framework to a young domain such as data visualization.

In Valcamonica, a valley close to Brescia in the north of Italy, there is the largest number of prehistoric petroglyphs in the world. Here, UNESCO identified about 140,000 different drawings. But the actual number is likely twice as much because some of them are still covered by vegetation. All these incisions date back to different ages: Epipaleolithic, Neolithic, Copper Age, etc., until the Middle Age. This corresponds to a long period, about six or eight millenniums, where people have used this kind of visual communication.

Historical information has been deduced from these drawings: people living in that area practiced agriculture, fought to protect their community, hunted wild animals, and prayed according to their religious beliefs. For thousands of years, people living there *represented* their world through visualization.

Today the scientific community refers to this practice as Information Design. Robert Jacobson, one of the pioneers in this field, defines Information Design as the discipline whose "purpose is the systematic arrangement and use of communication carriers, channels, and tokens to increase the understanding of those participating in a specific conversation or discourse". The conceptualization of this domain was first introduced in the 1970s and became official with the publication of the Information Design Journal in 1979. However, important thinkers such as Charles Joseph Minard, John Snow, Florence Nightingale and Otto Neurath previously carried out some significant works in this field.

In recent years, other areas of study entered Information Design with different denominations. One of these is Data Visualization, a recent domain that explores how digital data can be portrayed. Now "Data Visualization" as a term is in wide-spread use all over the world; it is common to come across writings, courses, and web sites related to this domain: FlowingData is one of them, a web magazine whose payoff is "Data Visualization, Infographics and Statistics".

This article expands on the notion that it can be reductive to only speak about visualization. In the past, people who lived in Valcamonica were not simply drawing what they saw; rather they used images to represent their community and their lives. What they drew was not just a sign, they also implied a behavior

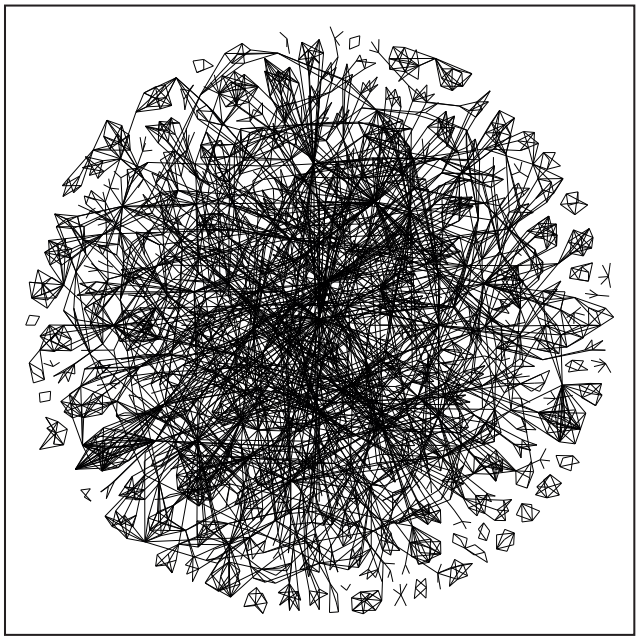

FIGURE 1: *The network visualization based on authors and keywords derived from publications.*

beyond that sign. Illustrations are meaningful because they represent something important to the community; consequently, it is fundamental that those who observe them also detect the object indicated so as to hear the voice of the community who drew the sign. To investigate this theme, the argument should be built by investigating the relationship between visualization and representation, as can be shown by a practical example of design; the brand image of DH2014, the Digital Humanities conference that took place at the EPFL and UNIL campus in Lausanne, Switzerland.

The idea was to represent the Digital Humanities (DH) domain as a pattern that could be beautiful and ductile, which would allow it to be used as a brand image for producing posters, covers, banners, etc. The DHLAB, laboratory in Digital Humanities at EPFL, one of the organizers of the conference, accomplished this task by using the conference data set—in particular the submission information. By analyzing this data it was possible to create a network visualization based on authors and keywords derived from the metadata found in all papers and posters accepted for the conference. All the keywords of each document were linked, as well as all authors of each document. Then, the authors and keywords of each document were linked. The three sets of links were merged to form a unique network that provided a representation of the DH community's complexity.

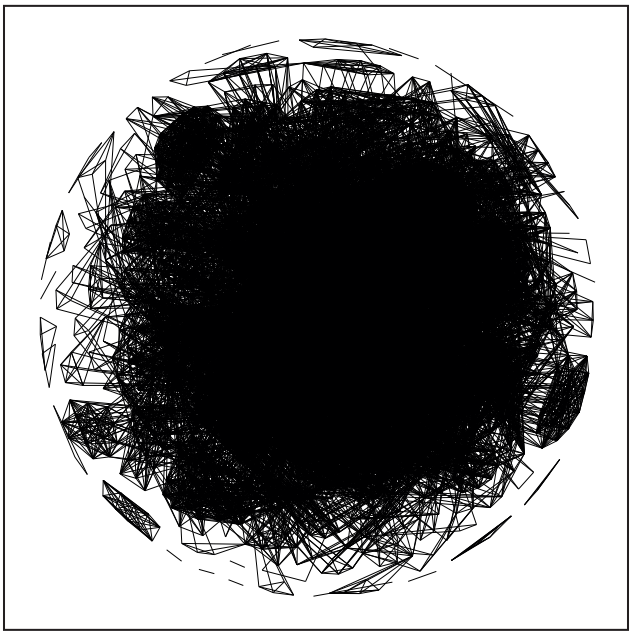

FIGURE 2: *Conference authors represented by co-authoring and shared keywords.*

Subsequently the original network was split in two networks: the first representing the authors, the second the keywords. The purpose was to *simplify* the visualization in order to make it more comprehensible.

This network represents all authors attending the conference who had entered at least one submission. The authors in the middle of the network are the most linked, both due to their co-authoring and to common keywords. In fact, this is not just a network showing who published with whom, but also a network displaying authors with shared keywords or, in other words, who worked on the same theme.

The force-directed graph, arranged by combining ForceAtlas 2 and Fruchterman–Reingold algorithms, makes identification of author clusters easy. Due to these algorithms, the spatial disposition doesn't have a disposition based on coordinates, rather its relevance is in terms of proximity; the closer two authors are, the more documents or interests they share.

The social network of authors was printed and placed in front of the conference's entrance. Due to its large size this visualization, reified in a carpet, gave participants a clear invitation to exploration. As shown in the photograph, authors were attempting to locate themselves on the map. What soon became a game was a perfect mix between entertainment and examination; each person followed their personal path within the social network.


FIGURE 3: *The authors network visualisation materialized in a red carpet, placed just in front of the conference entrance.*

Such a search generally led them first to spotting authors that were well known to them, then to finding their own colleagues, and finally themselves. Finding one's own name was a kind of success that triggered different behaviors, which were often shared on social networks as Twitter. Among the actions identified there were: a) a *portrait* when authors asked to have a picture taken of them, b) a *postcard* in case they found a friend or close colleague and sent them a message, c) an invitation to play the *game* when they invited other people to find themselves, or d) a *selfie*.

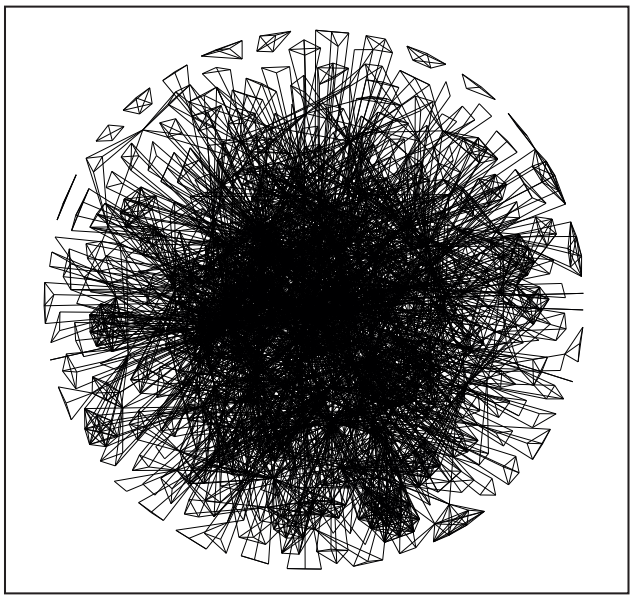

FIGURE 4: *Conference keywords represented in a network.*

This active interaction with the carpet was not mere engagement, since any form of data visualization can only be considered successful when it creates comprehension and knowledge among its viewers. Complex data visualizations require time to be understood; the aspects of entertainment, the exceptionality of the media, and social interaction involved at the 2014 Digital Humanities Conference made the process of understanding easier. Interaction with the carpet was not a solitary experience; it was a collective one where authors improved their comprehension of a form of describing a collective domain—the representation of Digital Humanities.

The network of keywords is probably the most interesting one. As the Digital Humanities community shows uncertainty in defining their very domain, this visualization is intended as a *representation* of the documents presented at the conference, of the authors attending the conference, of the conference itself, and, last but not least—of the domain of Digital Humanities.

The edges signify that two keywords are used in the same document, while the lines thickness is given accordingly to the occurrence of the connection. This thickness increases the depth of the layers—about twelve measures are used in the current network—thereby enhancing the reading with a sense of depth and highlighting the most used connections.

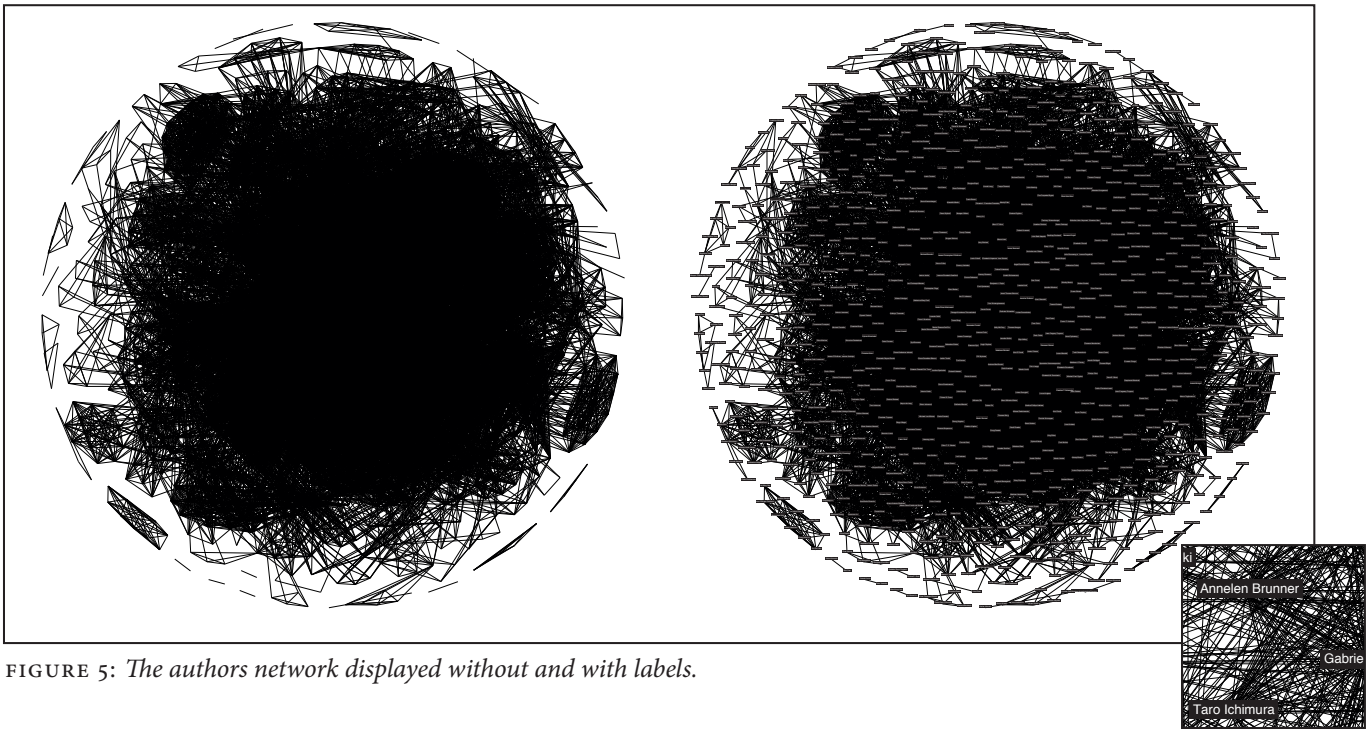


FIGURE 5: *The authors network displayed without and with labels.*

Charles Sanders Peirce has been a prolific mathematical logician and founder of American pragmatism. Peirce, with Ferdinand de Saussure and Charles W. Morris, was one of the most prominent theorists in Semiotics, and is famous for his contributions to the Sign's Theory, an approach based on the dyadic relationship between sign and object where the sign is something that can be interpreted, and the object is the target of the sign meaning. If a reader looks at the word "dog" in a book, automatically he transforms the word in the concept of dog that is, what is the word meaning in that context. In this example the sign is the word "dog" and the object is the concept of the dog, precisely the meaning of "dog" for the reader—that should be the common comprehension of the word dog.

Peirce, during his life, wrote a lot of definitions regarding Sign's Theory, such as the following: "I define a sign as anything which is so determined by something else, called its Object, and so determines an effect upon a person, which effect I call its interpretant, that the later is thereby mediately determined by the former."

Different from the others, according to Pierce, the theory of signs is not just based on a dyadic relationship formed by signs and objects, but also on the interpretant, a fundamental point of his approach which introduces an interpretation between the object and the sign. In the above example, the interpretant is the person read -ing the word "dog." Consequently the basic structure becomes a triple, which comprises the sign, the object, and the interpretant.

Peirce used to refer to the signifying element in different ways: sign, representamen or *representation*. Contrary to the meaning of the word "sign," "representation" bears a wider sense: while "sign" just refers to a visual element, "representation" encloses the sign and the object together.

Applying the Signs Theory to Information Design could be inspiring. Thanks to that theory, the authors' network of DH2014 can be interpreted in two ways: by assuming the network node is the sign, and the label—the nominal data associated to the node—is a sign extension, the object could consequently be 1) *the author*, whose interpretant is his written document, or 2) *the document*, whose interpretant is the author who wrote that document. By assuming the interpretant as the determinant of the sign/object relation, both versions are appropriate: 1) in the first case the document describes the relation between the node and the author, and 2) in the second case the author is the key to understanding that relation; by asserting his fatherhood, he takes on the responsibility to be associated with a certain scientific document. The act of authoring denotes the relationship sign/object.Both choices are reasonable, but by considering the keywords' network we will obtain further insights that will help identifying the right interpretation.

The interest of a second attempt rests in the meaning of keywords. In this visualization, nodes represent keywords, accurate words chosen by their respective authors appearing as the documents' metadata. The nodes

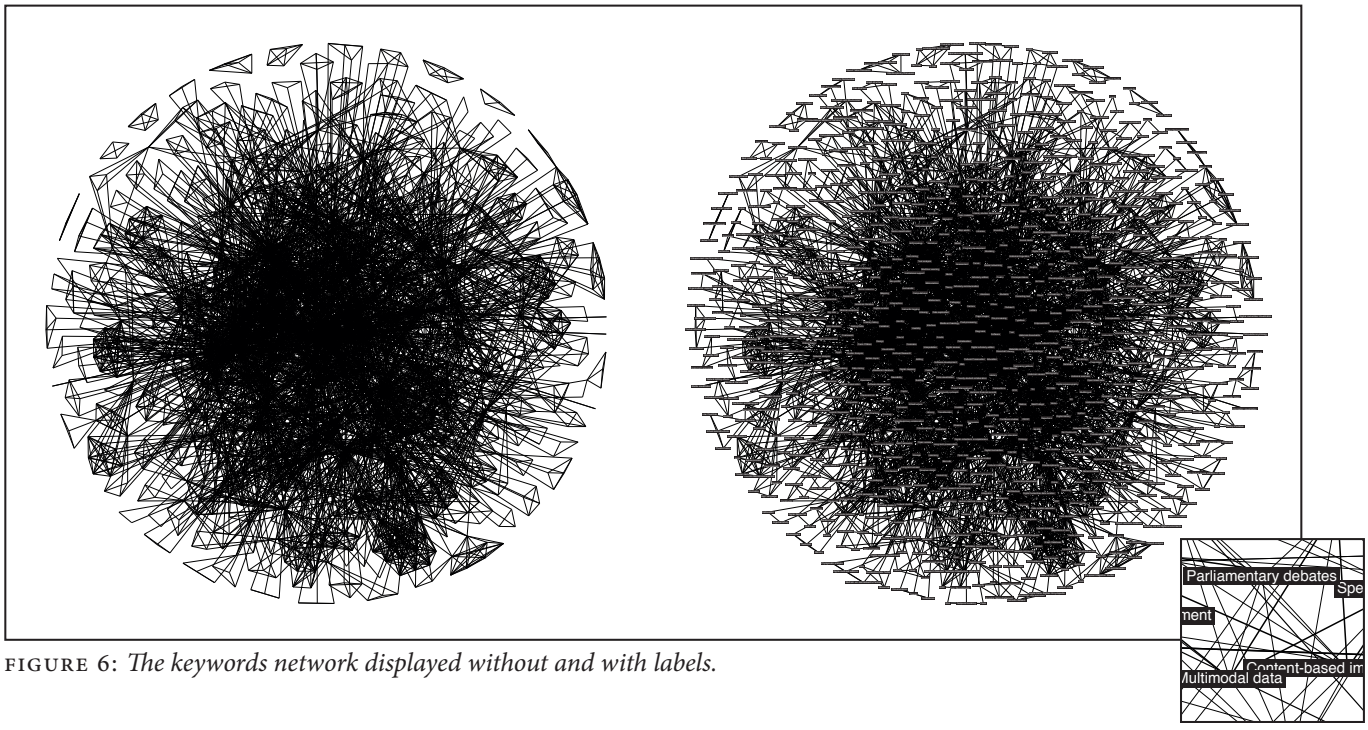


FIGURE 6: *The keywords network displayed without and with labels.*

are extended with nominal data, exactly as it was done for the authors' network. In this case, there are three possible ways in which it is plausible to apply Peirce's thought: 1) the object is the *meaning* of the keyword and the interpretant is the document, or 2) the object is the *use* of the term and the interpretant is the document—as an object authored by the writer—or 3) the object is the *document* and the interpretant is the meaning given by the author.

Evaluating the best interpretant is not an easy task, but the thinking of Ludwig Wittgenstein could be helpful to pursue the scope. Wittgenstein said: "For a large class of cases of the employment of the word meaning—though not for all—this way can be explained in this way: the meaning of a word is its use in the language." This statement suggests that scientific publications embody the specialists' language. In Philosophical Investigations where this statement has been extracted, Wittgenstein doesn't quote Peirce and, to be honest, Wittgenstein has never quoted Peirce in any document. However, Charles Sanders Peirce was such a prominent person, that everybody could agree on the point that Wittgenstein must have read Peirce. If considering the statement by Wittgenstein through the eyes of Peirce, "the meaning of the word is its use in the language" appears to be incredibly close to what Peirce defined as the interpretant. If he was to shift his attention to the visualization, Wittgenstein would have interpreted the keywords network in this way: a) the sign is the node with the nominal data, b) the object is the meaning of the word and c) the interpretant is what makes the relation sign/object understandable to the community: the use of the language indicates the meaning of a certain word, or simply the document intended as a *medium* of communication.

Considering this meditation on the Signs Theory, we can claim that data visualizations sometimes reveal a deeper meaning. Behind the visual apparatus, there is a projection that connects the visual part to something represented—a projection from signs to objects. Visualization has a reductive meaning when something is represented. In the DH2014 visualizations, the authors and keywords networks are specific representations of the Digital Humanities community in a particular moment. Behind the visual display, there is a real network composed of people and themes of research. The connection between their representation and the community's words is provided in the conference documents and the language used by professionals to describe their work.

In "From Realpolitik to Dingpolitik," the first text in "Making Things Public", Bruno Latour discusses the way of doing politics, but what is useful to the argumentation is how politics and the topics of interest in politics are represented in public spaces.

Latour describes his interpretation of the "object-oriented democracy" by bringing together two different meanings of the word representation: the first "designates the ways to gather the legitimate people around some issues;" the second "represents what is the object of concern to the eyes and ears of those who have been followed." It is possible to compare politics with a conference. For the DH2014 conference, one representation is given by all authors attending the meeting, which is the definition of Digital Humanities as a discipline.

If the comparison politicians/authors is explicit, the representation of the DH definition deserves a clarification: as Digital Humanities is quite a new domain, it is controversial to represent it, because of its diversity. As opposed to an authors network, the keywords network produces a special result, which is to assemble all of the documents' keywords into a lexical representation of the domain. This representation, albeit highly unstable in time, is a steady image of the DH community in the summer of 2014.

In the conference context, the object of interest is the definition of the community itself, a definition that was represented by means of a data visualization based on keywords. These keywords—the signs—are extended to the meaning of the words—the objects—whose understanding is given in the documents written by authors—the interpretants; this triple confers the visualization the authority of a representation.

Since the assembly is composed of the same authors who contributed to the conference, the keywords representation could be viewed as a loop, but it is not a loop reflecting the thoughts of each participant—the definition arises from the documents as a sum of voices, one for each author, and the object of concern is not a sum, rather it is a whole where each voice has the same dignity. Thus, an author, whose voice is part of the chorus, could disagree with a definition to which he has contributed.

To conclude, Latour asks, "How to represent, and through which medium, the sites where the people meet to discuss their matter of concern?" The answer is data visualization. As discussed, sometimes data visualizations could be better defined as visual representation because of what they represent. In the example of DH2014, data visualizations are designated as a representation of a community, of a definition, and of the central topic

of interest to be discussed at a meeting, and which can be criticized and modified following to the forces that drive the domain of Digital Humanities.

## BIOGRAPHY

*Dario Rodighiero* is PhD candidate at the Doctoral School of EPFL, attending the Doctoral Program Architecture and Sciences of the Cities. He is employed as designer at the College of Humanities in the DHLAB where his supervisor, Frédéric Kaplan, is director. Previously Dario joined the European Commission and the team of AIME, headed by Bruno Latour, at the médialab of Sciences Po. He created the brand design for the CHI2013 conference in Paris and for the DH2014 conference in Lausanne.

## BIBLIOGRAPHY

AA.VV. *Information design*. (MIT Press, 1999).

AA.VV. *Making things public: atmospheres of democracy.* (MIT Press ; ZKM/Center for Art and Media in Karlsruhe, 2005).

Atkin, A., *Peirce's Theory of Signs*, The Stanford Encyclopedia of Philosophy. Summer 2013 Edition.

King, B.A., Wertheimer, M., 2007. Max Wertheimer & Gestalt Theory. Transaction Publishers.

Meirelles, I., 2013. Design for Information. Rockport Publishers.

Purchase, H.C., Andrienko, N., Jankun-Kelly, T.J., Ward, M., 2008. Theoretical Foundations of Information Visualization, in: Kerren, A., Stasko, J.T., Fekete, J.-D., North, C. (Eds.), Information Visualization. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 46–64.

Rodighiero, D., 2014. Digital Humanities 2014: representing a controverted definition. Presented at the IC Research Day 2014, Lausanne. doi:10.13140/2.1.1420.1289

Wittgenstein, L., 2003. Philosophical investigations: the German text, with a revised English translation, 3rd ed. ed. Blackwell Pub, Malden, MA.