

Self-Recognition in Data Visualization.

How Individuals See Themselves in Visual Representations.

Par Dario Rodighiero et Loup Cellard

From Personal Identity to the Media.

« When we see any thing to be in any place in any instant of time, we are sure, (be it what it will) that it is that very thing, and not another, which at that same time exists in another place, how like and undistinguishable soever it may be in all other respects : And in this consists identity, when the Ideas it is attributed to vary not at all from what they were that moment, wherein we consider their former existence, and to which we compare the present »

— John Locke (1999, p. 328).

It was the 17th century when John Locke wrote about *personal identity*. In his eyes, it is a form of awareness that someone feels when a present experience recalls something from memory. This moment of identification takes shape between the present and the past, between what you are seeing and what you have already seen.

The concept of personal identity is intimately intertwined with the philosophical question of being an individual in the world. Who am I ? How am I different from others ? What makes me similar to others ? All of these personal questions introduce the key actors that take part in self-recognition : a) the *reader* who observes, b) the *subject* which the reader recognizes, c) the *group* to which the reader belongs, and d) the *representation* that recalls the subject. This list shows that the self-recognition does not only concern the reader but rather a network of human and non-human actors.

Readers play a central role in this network, being both subject and object of the visual representation. Their main characteristic stays in their unicity, which stems from a process of *subjectivation*. Michel Foucault refers to subjectivation as « the endeavor to know how and to what extent it might be possible to think differently, instead of legitimating what is already known » (Foucault 1985, p. 9). By saying that, Foucault stresses the intimate need of developing a distinct personality, which plays a major role in visual interpretation.

Subjectivation is a mutual force that operates in a space where the individual is opposed to the world. On the one hand, the individual is actualized by the world, and on the other hand, the world contains all the potential ways to design humans, like in a permanent opposition between actual and potential (Deleuze and Parnet 2007). This brings each individual to be shaped in a unique way that depends on the social environment, which plays a major role in subjectivation.

To make things more complicated, there is another level crowded by *mediators* between individual and the social context. These modify personal perception by operating on a threshold that Vilém Flusser figures out as a space of translation populated by the *media* (Flusser 2011). The process through which humans are shaped is, therefore, deeply influenced by the technology that translates the world for our senses. The importance of media was captured a few decades ago by Marshall McLuhan, who refers to them as the tools that shape humans after being shaped by humans themselves (Culkin 1967, 70). The subjectivation can take full advantage from media when we know and use them in an appropriate way (Kaplan 2014).

The Heideggerian human-world relations that guide our perception of the world are then influenced by technology, which acts as a mediator between humans and the world (Ihde 1990) (Verbeek 2018). This article focuses on a specific technology that influences our perception by translating data into images, data visualization.

Data Visualization for Representing Identities.

Data visualizations map textual and numeric information through personal computers. Although the art of visualizing information dates back centuries (Yates 1954), the digital turn of the 20th century marked a significant moment by welcoming the term « data visualization » in the universal language. Its rise takes place during the 1980s (see Figure 1) when Edward Tufte gathers a heterogeneous group of practitioners through a series of books (Tufte 2001).

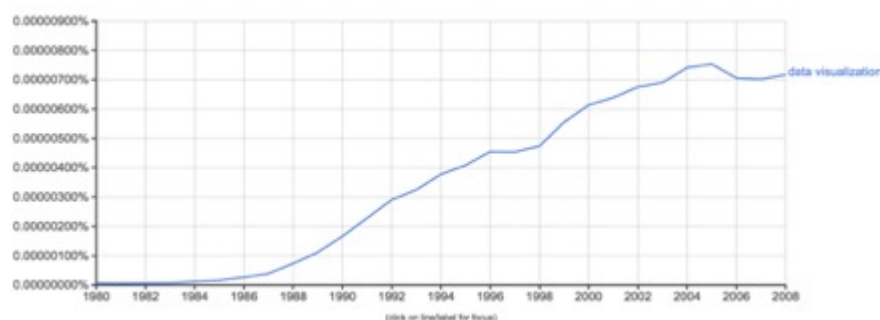


Figure 1. The Google N-Gram graph shows the frequency of the term « data visualization » over time. The increase of the 1980s points out the moment in which data visualizations become widely used.

Data visualizations come from a design process, which can be less or more elaborate according to the amount of information displayed. They are artifacts that result from a series of transformations and endeavors driven by a designer (Neurath and Kinross 2009). Visualizations make visible the information that otherwise would be impossible to see in accordance with the decision taken (Manovich 2008).

Although the data visualization literature is growing fast (Meirelles 2013), many of their features still remain unexplored. The article contributes to the analysis of the phenomenon of *self-recognition*, which happens when a reader recognizes the own digital identity in a data visualization. The process of recognition is explored from the beginning, when the digital self is made of by collecting information, to the end, when readers improves themselves as a result of a reflection after seeing their own identity visually represented.

The Course of Recognition.

This article is inspired by the last book^[1] written by Paul Ricœur, titled *The Course of Recognition*. The book begins by analyzing the meaning of the term recognition according to two French dictionaries, the *Dictionnaire* by Emile Littré and the *Grand Robert* (Ricœur 2005, p. 1-22). The book brings readers through a wide variety of definitions to finally narrow down to three of them, *recognition as identification*, *recognizing oneself*, and *mutual recognition* (David Carr 2008).

Applying these definitions to data visualization reveals some interesting mechanisms of the visualization reading. The *identification* establishes a relationship between the visualization and its readers. The *recognition* connects readers with their digital identity. The *mutual recognition* marks the reciprocal approval between readers appearing in the same visualization.

Two additional concepts extend the Ricœur's theory, the digital identity and the promise. The *digital identity* is the datafication of personal identity and the *promise* is the commitment that readers make after an intimate reflection for their future life^[2].

This article uses a case study displaying the EPFL Digital Humanities Laboratory headed by Prof. Frédéric Kaplan. The choice was guided by its multi-disciplinary structure that often

characterizes the domain of digital humanities, making this unit a perfect example to map. The laboratory data are transformed into a network displaying peers as nodes and collaborations as edges. This network is successively fragmented, revealing the collective evolution over time (see Figure 2).

On the horizontal axis, each network shows collaboration using edges based on publication authorship ; on the vertical axis, a thread between networks, called *trajectory*, stays for the presence of someone for two consecutive years. The idea behind this representation is to disclose the laboratory history using social ties and time continuity (Rigal and Rodighiero 2015) (Rigal, Rodighiero and Cellard 2016) (Rigal and Rodighiero 2017).

Using this case study to support the argumentation, this section presents the five concepts of self-recognition in this order : a) the *digital identity* of laboratory members, b) the *identification* with data visualization, c) the *self-recognition* in digital representation, d) the *mutual recognition* between members, and e) the *promise* for future choices.

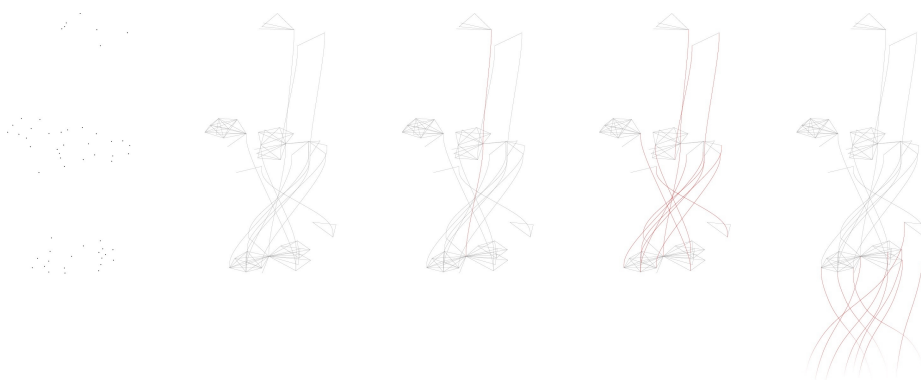


Figure 2. This series of images represents the EPFL Digital Humanities Laboratory over time and through collaborations. The same network visualization is declined in five configurations, which support the article's argumentation on self-recognition in the next sections, 1) digital identity, 2) identification, 3) self-recognition, 4) mutual recognition, and 5) promise.

Digital Identity

The first phase of self-recognition happens in information systems, whose main task is keeping safe computer records and serve them on demand. Although they could appear as functional

storage technologies, information systems are non-human agents intertwined part of the social environment. They can certify information in society, for example, by labelling the status of individuals. This characteristic makes them a sort of technical authority.

If information systems can certify personal information, it is more than legitimate to wonder who controls them. This question leads us to think about the complexity of digital authorship (Keen 1981, p. 28). An information system is usually maintained and updated by one organization, which is stirred by a management team through decision-making. This management implements policies that regulate data in terms of heterogeneity, accessibility, and ethics. A government, for example, certifies citizenship by saving a record in a database, which extends the capacity of the federal organism.

Large organizations, such as the EPFL, distribute data concerning their employees in different systems. Each employee is associated with a unique *identifier* that tags personal records. Contracts, founding, formal partnerships, and publications are always associated with one or more identifiers. Then, records are retrieved using identifiers as access keys for digital identities.

Employees are quantified through *datafication*, which translates daily activities into data. Datafication is constantly augmenting due to devices such as identity cards, mobile phones, laptops, and smartwatches^[3]. These devices guarantee the traceability of individuals by producing digital traces, which can be later reassembled to retrieve personal data in digital storytelling.

Ulysses, in the *Odyssey*, offers a beautiful example of storytelling when returning to his land ; he asks Ithaca citizens to be recognized as the king by addressing a public speech (Ricœur 2005, p. 72). The identity was oral at that time as demonstrated by his speech, but, today, identity is not only written down but also digitalized. In this regard, Jos de Mul introduces the notion of *database identity* by arguing « that databases [...] transform the narrative identity of pre-modern and modern persons and cultures into what we might call a database identity » (De Mul 2015, p. 98). Today, information systems represent the foundation of self-recognition, keeping the *moral accountability* fed daily by our devices (Ricœur 2005, p. 106). If Ulysses reclaimed his identity as a king by telling his story, people today can establish their identities through personal data. The information system certifies our identities and behaviors.

EPFL owns an information system that certifies academic literature called Infoscience. Authors are in charge of uploading and keeping records updated. Taking care of personal records is a task that guarantees visibility on the scientific community. Figure 3 shows how DHLAB members are

extracted from the database and mapped on the canvas. It is easy to notice how a visualization showing unique identifiers without relational information might be meaningless.

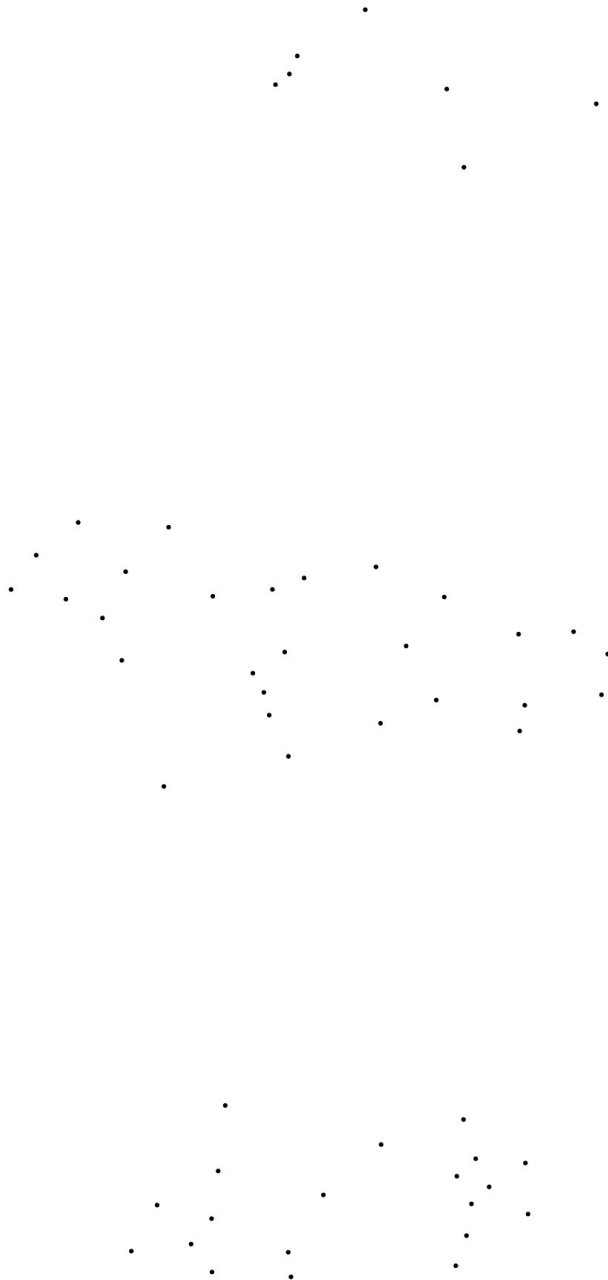


Figure 3. Digital identities of DHLAB members are recorded, but invisible without projection. Mapping them without any relational information shows them as a bunch of unorganized nodes.

Identification.

Transforming data into visualizations is the result of a design process. Database records are not ready to use as data need to be collected, selected, cleaned, normalized, and checked in terms of quality and integrity. The team in charge of this process needs to make a treatment on raw data, which to some extent might recall to an ethnographic study. If ethnographers collect data by observing specific behaviors on fieldwork, the designer's task is about checking the integrity between available data and observed or supposed actions. While ethnographers create data, designers verify data. Primarily when they are employed, designers investigate an organization like an ethnographer would work on the fieldwork. Data are collected from information systems as well from people by talking with them, which makes the designer's work both technical and social.

Although the discipline of design studies employs methods from ethnology to « determine design requirements or iteratively refine a design » (Munzner 2008, p. 139), there is a kind of ethnography that looks at the consistency of information. Designers should always validate the accuracy and origin of data as these affect the final result. Verifying information quality is a professional obligation. Therefore, the designer's work does not consist of creating beautiful visualizations, but rather of applying ethnographic methods to the design practice (Salvador, Bell and Anderson 1999) (Barab 2004) (Nova et al. 2015). Whereas Bruno Latour and Steve Woolgar made the observation an important feature for studying scientists in laboratories (Latour and Woolgar 1986), designers use observation to check information accuracy of translated data into visual forms in the most respectful way possible. The designer's ethic signifies to « formulate and support morally good solutions » (Floridi and Taddeo 2016, p. 1) because « trust and transparency are crucial » (Floridi and Taddeo 2016, p. 3). Understanding data and appropriate design are at the heart of the designer's work to avoid any kind of bias and misunderstanding. Paraphrasing Steve Woolgar and his concept of ethnographic inversion for the history of science (Woolgar 1989, p. 202-203), visualization is not something to take as granted because it might contain errors ; an examination is needed to prevent any form of misleading information from the very beginning of the design process.

« The ethnographic stance insists that representation is not something to be taken for granted when studying science. It is instead an aspect of the culture of science when requires examination. The corollary, of course, is that it is inappropriate to adopt the notion of representation in one's own study of science. It is a topic rather than a resource or, more exactly, representation is both topic and resource rather than just a resource. To adopt or take for granted a notion so central to scientific practice would be a direct disavowal of ethnographic inversion ; it would be to go native before the study had begun »

– Steve Woolgar (1989, p. 203).

Thus, the information system embodies an already existing representation, which makes data an object *given* to the designers. Paul Ricœur refers to Immanuel Kant for introducing such the concept of a *priori* (Ricœur 2005, p. 38). Kant states : « Now if sensibility were to contain a *priori* representations, which constitute the condition under which objects are given to us, it will belong to transcendental philosophy » (Kant 1998, p. 152). Kant's words argue that the visualization might be based on the pre-existence of information. However, a *priori* data has to be validated by the designer. Johanna Drucker takes up the thesis of Bruno Latour and argues that information is not something *given* as the Latin etymology suggests instead taken, saying that *capta* is a more appropriate term for data that are collected (Drucker 2011, §3). Collecting underlines a decision-making process that is part of the design (Van Es, Lopez Coombs and Boeschoten 2017). Visualizations are therefore the result of a performative act determined by a designer situated in a precise cultural and social context (Drucker 2011, §29).

In the case study, the designer draws a laboratory from a *priori* data from the Infoscience. The DHLAB members recognize the system's authority, which gives credibility to the visualization. Figure 4 illustrates the laboratory's collaborations where each and every network displays the peers linked using publications, from 2013 at the top, down to 2015 at the bottom. Vertically, the visualization connects the same individuals over time using the trajectories.

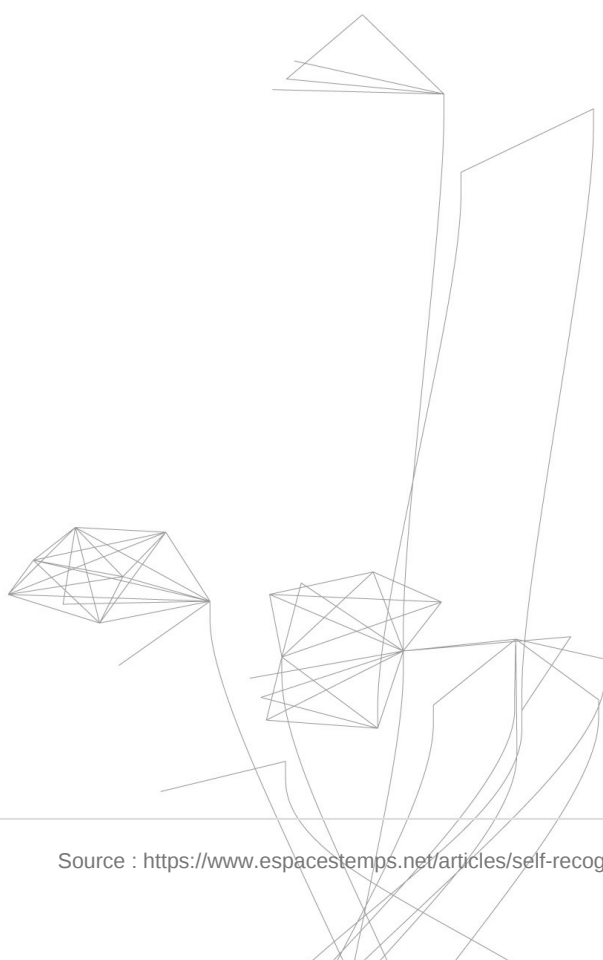


Figure 4. Digital identities of DHLAB take form in a relational space. Publication co-authoring forms network edges and nodes stay for individuals. This visualization aims to stress the laboratory temporality using long vertical trajectories (Rigal and Rodighiero 2015) (Rigal, Rodighiero and Cellard 2016) (Rigal and Rodighiero 2017).

Self-Recognition.

Recognition is a process that takes place at different scales. If recognizing the visualization's authority has to do with the large scale, focusing on single elements means working at the details.

For René Descartes and Immanuel Kant, recognition is an act of knowing to identify a single unit of meaning above all, but they pursue it differently. Descartes employs *differentiation* to discriminate one thing from another ; Kant makes use of *association* to stress the similarity (Ricœur 2005, p. 27-28). No matter what kind of approach is chosen, both help in establishing a relationship between elements that readers use to decode visualizations.

Reading is a performative act *within the context of the intentionality of consciousness* (Stiegler 2008, p. 61) aimed to understand visualizations *in a particular set of circumstances* (Drucker 2011, §30). If design is about representation, reading is about the observation of patterns that bring to a new perception of existing information. Reading is a way to amplify knowledge through a personal reflection and a collective discussion in a logic of constructivist epistemology (Glaserfeld 2014). Its subjectivity makes each act of reading unique and personal. W. J. T. Mitchell refers to the visual understanding as the mental eye or the intimate view (Mitchell 1995, p. 51). Reading can also be seen as a close relationship between the visualization and its reader. If Marcel Duchamp claims there is no essence in art without the viewer (Duchamp 1994), likewise no visualization bears meaning without readers as there would not be any form of recognition.

Reading has a specific configuration that we call *self-recognition*. Seeing insights can bring readers to recognize themselves by recalling memories associated to their identity. This means that visualizations can embed familiar entities that establish a closer relationship. This sense of recognition, already discussed in the context of social networks (Cardon 2008) (Georges 2009), remains unexplored for visualizations.

Visualizations can work as an optical instrument connecting readers' memory with information. If reading can be defined as the whole deed of understanding, the self-recognition occurs exactly in the moment when the reader says *it's me* (Ricœur 2005, p. 254). Only in that moment readers establish a close tie with their own selves. If the visualization is a translation from society through data, the self-recognition seals the loop by doing the inverse pathway (Rodighiero and Romele 2020).

This seal finds an analogy *In Search of Lost Time* that Ricœur takes as an example of the *proof of veracity* (Ricœur 2005, p. 68). When readers recognize themselves, the evidence of truth is achieved. Self-recognition is the key to validate a visualization that bears information concerning the reader. Data visualization works as the Proust's optical instrument, which readers use to see deeper into their own self.

« In reality every reader is, while he is reading, the reader of her own self. The writer's work is merely a kind of optical instrument which he offers to the reader to enable him to discern what, without this book, he would perhaps never have perceived in himself. And the recognition by the reader in his own self of what the book says is the proof of its veracity, the contrary being also true, at least to a certain extent, for the difference between the two texts may sometimes be imputed less to the author than to the reader »

— Marcel Proust (1993, p. 218).

Figure 5 shows in red a specific trajectory that represents one member of the DHLAB over time. Readers establish a connection with the visualization in which they can recognize their own self. All of the graphical elements employed stay for the data stored in the Infoscience system. When represented information matches readers' memories, the data visualization passes the test by confirming its proper authority.

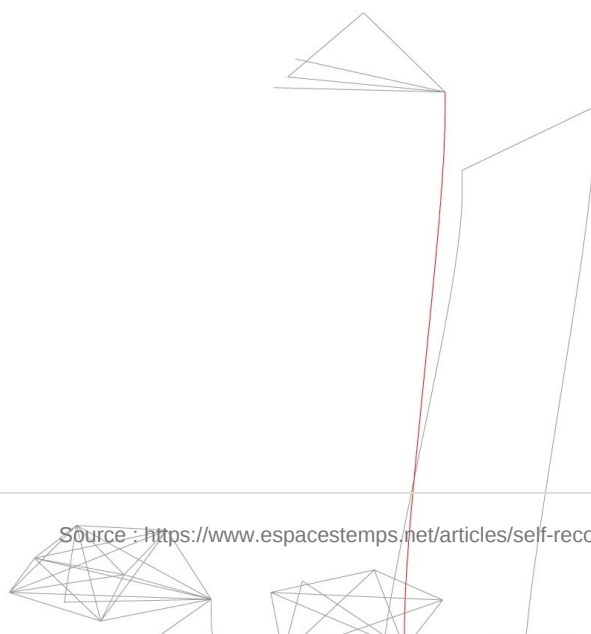


Figure 5. Self-recognition happens at this stage individually, when a reader recognizes the own identity in the red trajectory. Personal memories are strengthened during the performative act.

Mutual Recognition.

At the end of *In Search of Lost Time*, the narrator attends a dinner organized by the Prince de Guermantes. During this dinner, he encounters the past by meeting the acquaintances and friends of his life. He tries to recognize each of them in a timeless narrative of his life, but they are difficult to recall due to the time (Ricoeur 2005, p. 150-151). The narrator recalls the events that characterized his life in the act of recognition, shifting the focus from himself to the *whole group*. If the proof of truth previously was about the identification with the self, this further form asks for a mutual recognition performed by all the members of the group.

Readers realize that self-recognition is not individual performance, but rather a collective act interpreted simultaneously. Reading brings together all the readers in front of a general memory, which is different from the sum of all the single memories. Individually, the mutual recognition entails an act of self-recognition first, then the readers move the glance towards the other individuals that appear in the visualization. This is the moment in which the reader is aware of being part of a group by recognizing the identities of other members. When this act is multiplied for all of the readers, a complex system of relationships happens not only between the readers and the visualizations but also between the readers. This specific relationship between the readers makes the recognition *mutual* ; when readers look at each other, they recognize and are recognized. When Ricoeur refers to mutual recognition, his writing becomes intimate, making the encounter with acquaintances and friends the most touching moment of personal identity. He writes, « it is indeed our most authentic identity, the one that makes us who we are, that demands to be recognized » (Ricoeur 2005, p. 21). As Gloria Origgi writes, the social self is an image given by the deviation and multiplication from the gaze of others (Origgi 2016, p. 3). We recognize the others, and we ask others to accept us to be part of social life.

The different dimension of the own self that has been discussed so far roughly corresponds to the classification introduced by William James. For him, the self can be material, social, and spiritual, assuming its most abstract representation in the pure ego (James 1890). a) The material self is the physical part of the individual, such as the body, the hairstyle, or any of the things we use daily. b) The social self is described as the totality of the different attitudes a person assumes according to the social environment as we act differently at work and at home. c) The spiritual self

is related to the reflective process of the individual ; in other words it is about thinking of ourselves as thinkers. d) The pure ego is the complete abstraction of the first principle of subjectivity in each of us.

Concerning James's ideas, we depict the self as a radial emanation that surrounds the body in a sequence of membranes, one inside the other. This emanation encompasses different spaces which are structured like spherical spaces. The thickness of each layer changes continuously according to the social environment. While innermost layers make up the thinking part that is at the heart of the identity recognition, outermost layers are the external ramifications that make interactions with other people possible. Whereas the inner layers are intimate, the outer layers represent the exteriorization of the self (Cardon 2008, p. 98). This exteriorization situates the individual into the *setting*, a notion that Erving Goffman described as the personal front in a theatrical metaphor of life (Goffman 1956, p. 14). These membranes reflect the structure of self-recognition in data visualization, in which the readers first recognize their own self and later the group. In this context, data visualization plays the role of the setting where the recognition takes place.

Figure 6 highlights the trajectories of the laboratory members. Each one recognizes not only their own selves but also other peers. Recognition is a basic behavior of society and academic life is not excluded. When someone joins a research group, the leader has to be recognized as well as its members, and all of them have to accept the newcomer. The mutual recognition is also a mark of esteem without which no constructive cooperation is possible. Society is based on a spirit of collectiveness within groups and of distinction between them. When a laboratory such as the one of Digital Humanities presents its research during conferences, the recognition is crucial for the image of the professor of the collective as well. The visualization, in this sense, helps to represent the group as sole entity. Sharing visual representations like Figure 6 is a way to create a sense of collectiveness between laboratory members and establishing the joint identity of a group (Rodighiero 2018). The collective digital identity thus becomes a way to stimulate a debate in and out of the group itself.

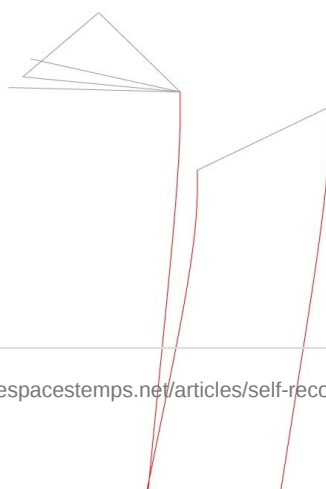


Figure 6. Each member of the laboratory recognizes his peers. All the red lines are not seen any longer as separate elements, but rather as elements of a whole. This operation corresponds to a shared sense of collectiveness.

Promise.

Marcel Proust always writes in a same pattern, projecting his own self a few years forward in an attempt to chase his own future. This effort results in making decisions for the future (Ricœur 2005, p. 109), so that the individual « finds in the capacity to make promises the criterion of its ultimate difference from identity as sameness » (Ricœur 2005, p. 103). In this sense, the visualization works as a time machine able to put the reader in the condition of envisioning the future, resituating the digital identity in the right place.

This capacity of envisioning the future emerges from the reader's understanding acquired by elaborating on the visualization. New knowledge emerges from the visual representation of digital identities. Self-recognition is an event, from which readers subjectivate by looking at their past. This process can appear as personal storytelling or a self-retelling in which the reader is the main character. The visual identity becomes a narrative identity through self-recognition, and the readers look at the future as individuals as well as a group.

In the DHLAB, the laboratory members look at themselves, stimulating by the collective representation. Where can we improve ? What is the next step ? How can we collaborate to be more effective ? The laboratory composed of individual members reflects upon its own form and makes itself a promise for the future : increasing interdisciplinarity, identifying new collaborations, incrementing publications, and developing creativity. These are some of the promises that might be made to project the collective towards the future. Figure 7 shows how the trajectories predict the future to forecast and control the evolution of the collective.



Figure 7. The laboratory starts to get a sense of what the next step will be and imagines the trajectories of all the members extended over time. The data visualization is an instrument for self-reflection to plan the future.

« There is no making of oneself [...] outside of a mode of subjectivation [...] and, hence, no self-making outside of the norms that orchestrate the possible forms of a subject may take. [...] The very being of the self is dependent, not just on the existence of the other in its singularity (as Levinas would have it), but also on the social dimension of normativity that governs the scene of

recognition »

— *Judith Butler (2005, p. 17).*

The process of self-recognition changes perspective about norms. No social recognition exists without a set of rules, which today rely on data. The performance of the DHLAB, for example, is measured using data such as scholarly publications, grants application, and teaching evaluation.

The relationship between data and norms is established in the academic community as more in general in large organizations. Records stored in information systems are regulated by norms set by organizations. EPFL, for example, measures the performance of the DHLAB members through different information systems such as Infoscience and IS-Academia. Yet it is crucial to think that creating an information system is the result of decision-making.

In this process, the designer has limited power as visualizations heavily rely on data. However, the design process is not only a one-way translation from data. The designers, now more than ever an ethnologist on fieldwork, can point out some issues to the decision-making body. There are multiple reasons why a design process can point out a lack of quality, such as the inaccuracy in publication records (Rodighiero, Kaplan and Beaude 2018). Confronted with the evidence, the management can update the rules or keep the current ones. The role of designers stops with the attempt to make explicit the non-visible norms.

Likewise, the reader is also an active actor in the design process. Readers can make use of visualizations as a Proustian instrument by retelling their past to plan the future. Inferring the norms from reading, they can change their habits and, consequently, their records. They can play or not this game within the organization ; all individuals act differently. However, in the daily activities of self-design (Groys 2008), it is important to be aware of the mechanisms that rule our digital identities and how in which extent we can affect the decision-making process. In academia, for example, conducting research changed drastically by the introduction of h-index. Today, scholars, especially ones looking for tenure, put a lot of effort in communication to have the highest number of citations. In this sense, Albert-Laszlo Barabási is a great example of someone that broke the rules of the game by collecting more than Nobel prizes.

To conclude, the designer plays two unexpected roles : the ethnologist who studies the social environment in the fieldwork and the archivist who pulls hidden information from the subsoil. Designer's role epitomizes a situation where « multiple methodological traditions intersect in digital devices and research » (Marres and Gerlitz 2016). There, the tools used by the designer

play the role of an operator interfacing an ethnography with archive excavation to frame individual and collective identity.

If Proust were still alive, he would invite readers making the most of data visualizations, saying : « Look for yourself, and try whether you see best with this lens, or that one, or this other one » (Proust 1993, p. 322).

Note.

This article was initially published as a manuscript in 2016. As we think it is still worth working on, we developed the original version for a journal publication. Since the first structure has been maintained, we kept the original title modifying the subtitle only. This version has been updated thanks to the comments of friends and colleagues, in particular, Nadine Baumer, Chloe Moon, Alexandre Rigal, Alberto Romele, and Tommaso Venturini.

This article is part of grant Early Postdoc.Mobility P2ELP1_181930 *Worldwide Map of Research* funded by the Swiss National Science Foundation (SNSF).