Universidade Federal do Rio de Janeiro
Escola de Comunicação
Programa de Pós-Graduação em Comunicação e Cultura
Linha de Pesquisa Novas Tecnologias e Estética

Marina Pantoja Boechat

To visualize, to discover and to share

On the uses of information visualization for building shared spaces for public debate, the cases of data journalism and controversy mapping

Rio de Janeiro

2016

Marina Pantoja Boechat

To visualize, to discover and to share

On the uses of information visualization for building shared spaces for public debate, the cases of data journalism and controversy mapping

Tese apresentada como requisito para obtenção do grau de Doutor em Comunicação e Cultura, no Programa de Pós-Graduação em Comunicação e Cultura da Escola de Comunicação da UFRJ, linha de pesquisa Novas Tecnologias e Estética.

Orientador: André de Souza Parente, ECo-UFRJ

Co-orientador: Tommaso Venturini, Sciences Po Paris

Rio de Janeiro

2016

CIP - Catalogação na Publicação

B669v

Boechat, Marina Pantoja
Visualizar, descobrir e compartilhar: Sobre os usos da visualização de informação para construir espaços compartilhados para o debate, os casos do jornalismo de dados e da cartografia de controvérsias / Marina Pantoja Boechat. -- Rio de Janeiro, 2015.
277 f.

Orientador: André de Souza Parente.
Coorientador: Tommaso Venturini.
Tese (doutorado) - Universidade Federal do Rio de Janeiro, Escola da Comunicação, Programa de Pós Graduação em Comunicação, 2015.

1. visualização de informação. 2. debate público. 3. cartografia de controvérsias. 4. jornalismo de dados. I. Parente, André de Souza, orient. II. Venturini, Tommaso, coorient. III. Título.

Elaborado pelo Sistema de Geração Automática da UFRJ com os dados fornecidos pelo(a) autor(a).

Marina Pantoja Boechat

To visualize, to discover and to share

On the uses of information visualization for building shared spaces for public debate, the cases of data journalism and controversy mapping

Tese apresentada como requisito para obtenção do grau de Doutor em Comunicação e Cultura, no Programa de Pós-Graduação em Comunicação e Cultura da Escola de Comunicação da UFRJ, linha de pesquisa Novas Tecnologias e Estética.

Aprovada em 29/04/2015

André de Souza Parente, doutor, ECo-UFRJ (orientador)	Tatiana Marins Roque, doutora, Instituto de Matemática da UFRJ
Marcio Tavares d'Amaral, doutor, ECo-UFRJ	Marco André Feldman Schneider, doutor, IBICT/UFF
Henrique Antoun, doutor, ECo-UFRJ	Fernando Álvares Salis, doutor, ECo-UFRJ
Ivan Capeler, doutor, ECo-UFRJ (Suplente)	Gustavo Silva Saldanha, doutor, IBICT/UniRio (Suplente)

Abstract

The thesis seeks to understand if in societies with ubiquitous data, that are heavily mediated by digital technologies, visualization can collaborate to compose shared spaces for public debate. This would involve making complex aspects of social life and public agenda visible. For this, first it was important to investigate the roots of visualization and its joint development with data, which is increasingly used as a basic unit to describe, record and manage social interactions. Second, it was also necessary to discuss the practices of information visualization and the challenges for its use as a tool for equipping public debate, especially in the fields of media and social sciences, where the issues of representing social reality are crucial. Therefore, the thesis was organized in two main approaches: first, we composed a small history of what we see as a process of discretization and reassemblage of supports and forms of information. With that, we seek to highlight the relation of mutual constitution between data and visualization, de-naturalizing the ubiquity of the former and contextualizing the importance of the latter. Second, we carried empirical work, by means of direct observation and qualitative interviews in two fields: data journalism and digital methods of social research, more specifically controversy mapping. Our main findings are mainly towards establishing the importance of the many and progressive transformations visualizations go through between visual analysis and final presentations, therefore considering visualizations as traces of debates and many interpretations along the way; and also the different perspectives and ways in which visualizations are accessed, organized and related at each moment. Finally, it led us towards the understanding that these practices point to a reversal of some traditional informational flows, by generating and reprocessing data that comes from digital tools into new mediation devices.

Keywords: information visualization, public debate, controversy mapping, data journalism

Marina Pantoja Boechat

Visualizar, descobrir e compartilhar

Sobre os usos da visualização de informação para construir espaços compartilhados para o debate, os casos do jornalismo de dados e da cartografia de controvérsias

Resumo

Discutimos como, em sociedades pesadamente mediadas por tecnologias da comunicação e da informação, onde os dados se tornam ubíquos, a visualização pode colaborar para compor espaços comuns para o debate público, o que envolveria tornar visíveis aspectos complexos da vida social e da agenda pública. Para tal, entendemos que é importante investigar as raízes da visualização e seu desenvolvimento em conjunto com os dados, cada vez mais utilizados como unidade básica para descrever, registrar e gerir as interações sociais. Em segundo lugar, é necessário também discutir as práticas de visualização e os desafios para seu uso como ferramenta para instrumentar o debate público, especialmente nos campos da mídia e das ciências sociais, onde as questões relativas à representação da realidade social são cruciais. Para tal, a tese foi organizada em duas abordagens principais: primeiramente, foi composta uma pequena história do que entendemos como um processo de discretização e recomposição de suportes e formas de registro. Procuramos com isso ressaltar o vínculo de mútua constituição entre a visualização e os dados, desnaturalizando a ubiquidade destes últimos e contextualizando a importância dos primeiros. Em segundo lugar, realizamos um trabalho empírico, por meio de observação direta e entrevistas qualitativas, em dois campos: o jornalismo de dados, e os métodos digitais de pesquisa social, mais especificamente a Cartografia de Controvérsias. Nossos principais achados vão rumo, primeiramente, à percepção da importância das progressivas transformações nas visualizações entre a análise e os resultados finais – consideradas por nós como rastros do debate e das diversas interpretações ao longo do caminho –, e também das diferentes perspectivas e maneiras pelas quais as visualizações são acessadas e relacionadas a cada momento. Finalmente, rumo a uma compreensão de que tais práticas apontam para uma reversão de alguns fluxos tradicionais de informação, por gerarem e reprocessarem dados a partir da informação circulante, utilizando-os como fontes para novos aparelhos de mediação.

Palavras-chave: visualização de informação, debate público, cartografia de controvérsias, jornalismo de dados.

Cette thèse est dedié à Ricardo. Pour tes blagues et pour tes yeux; pour tous nos jours, notre partage et notre amour.

Acknowledgements

Only now, when the work is almost over, I realize that this thesis was a huge journey. This text, with all its flaws and limitations, would not be possible without the outstanding people I met along the way, either just for brief moments or longer exchanges. I do hope it shows the traces of these encounters, disputes and collaborations. Listing names will probably make me forget some very important people, but I will take that chance.

First, of course, I thank all the professors at ECo-UFRJ, for the inspiration and for giving me a rare and precious view of the field. André Parente, Marcio Tavares d'Amaral, Henrique Antoun, Fernando Salis, Fernanda Bruno, Janice Caiafa, Paulo Vaz,

Tommaso Ventiruni, I cannot thank you enough for your generous input, for having me at the médialab and providing me with many valuable tools to face a huge challenge. Also all the amazing and equally generous guys from the médialab and Project Emaps: Anders Munk, Audrey Baneyx, Axel Meunier, Débora Pereira, Guillaume, Mathieu Jacomy, Richard Rogers, Peter Gerry, Liliana Bounegru,

To all the outstanding professionals who generously granted me interviews and allowed me into their work places. Especially Mario Leite and all the guys from GloboEsporte.com, André Hippert, Mario Kano and everyone from Folha de São Paulo, Rodrigo Burgarelli, Daniel Bramatti and the fantastic team at Estadão Dados, Mar Cabra, Rubens Paiva and Gabriela Allegro at Jornal O Globo, Gabriel Gianordolli, Marcelo Stoppa and many others.

To new and old friends, the ones that were not cited yet, even if most of them do not realize how much they are a part of this. Because, as the brazilian saying goes, in friends you have got everything. Thank you Marco Schneider, Rafael Cardoso, Patricia Breves, Rafael Vaz, Ana Ash, Aline Mathias, Daniela Fernandes, Priscila Vieira, Marcus, Camille Audouy, Simona Dantone, Raphael Vasconcellos, Anna Letícia, Helle Patricia, André Malheiro, Ana Paula Costa, Aline Leal, Mariana Orioli, Mayanne Linz, Gustavo Autran, Ana Carina Santos, Daniel Bezerra, Juana Huamán Charret, Lilian Martins, Ricardo Pimenta, Gustavo Saldanha… the list goes on.

To family. Extended. With much love.

Table of Figures

Table of Contents

1. INTRODUCTION

In societies that are heavily mediated by digital tools, more and more of our experiences, knowledges and practices is described as data, in order to be integrated in information and communication systems. This collaborates to a deluge of data and information that has been raising much concern on to how to store, recover and understand all these informational flows and reserves. Nevertheless, historically we have had other informational *booms*: many authors will indicate the invention of the movable type and the press as a turning point to information accumulation, that triggered a host of other devices for the production, organization, circulation and even political control of texts, such as libraries, archives, scholarly disciplines etc. (see, for example, CHARTIER: 1998). Others will point to the origin of the codex and even to the invention of phonetic writing (see, for example, BOTTÉRO:1995). To an extreme, studies in cognitive science will show that visual perception itself it coupled with procedures that reduce the amount of information that will be registered, in order to avoid overloads and better manage the available resources (see, for example, KIRSH, 2000, 2006, 2010a and 2010b; WARE, 2003 and GIBSON, 2014). In order to differentiate the information overflow that characterized the last decades of the twentieth century onwards, we could, of course, argue in terms of quantity and talk about the many terabytes of information produced and the amounts that increase geometrically (see, for example, WURMAN, 2000). But that would only describe information from a computational point of view, that is of course the description that best matches the *ethos* of our computational age and a specific kind of information. So we should focus on the specific character of our informational boom: the fact that it is data-based and works from inside digital technologies. This presents at the same time specific challenges and specific opportunities.

We have, on the one hand, a multiplication of spaces of discourse in the varied information architectures constructed in information systems, each one with specific rules, that select and aggregate messages to be circulated in specific ways, and not always will be compatible with one another. In this sense the many discursive spaces that have always existed in societies will be multiplied and reified in technical devices that often work to naturalize many discursive strategies. On the other hand, due to a wider variety of sources of recorded data and information, to the fragmentation of messages and to the discrete nature of data, information gets scattered as much as it gets ubiquitous, breaking contexts of reference. The challenge of achieving inclusive representations of social life, that has existed in many different ways

throughout human history, becomes fundamentally connected to the related tasks of overcoming platform biases in technical devices and to reassembling context from scattered references.

In this sense, we see the visualization of data and information as a very important resource for dealing with the present informational landscape. It has quite nebulous origins that are mixed with the origins of cartography, but made its first major advances at the dawn of statistics from the eighteenth century on, connected with the needs of the management of accumulated records, like in the work of Priestley, or public and commercial administration, a bit later with Playfair (ROSENBERG & GRAFTON, 2010). From then on, and especially with digital technologies, visualization has penetrated the most varied spheres of experience in interfaces, digital tools and others. It has become a tool for describing the informational flows with which we deal on a daily basis and also an interface for access to information and for interaction with many tools.

On the other hand, in the last few years, the use of visual data analysis – and therefore exploratory visualization – in scientific research in the humanities and social sciences has dramatically increased. In an environment of ubiquitous technical mediation by digital tools, every interaction generates more traces, which collaborates to a deluge of data. The challenge of making sense of this affluence becomes ever more important because, with these records and the growing formalization of our knowledge and exchanges in the form of data-based content, many aspects of our reality can now be tracked, treated and made visible (see VENTURINI & LATOUR, 2010; Rogers, 2013). Visualization comes in as an important tool throughout the whole scientific process when data is used: for the data analysis, but also for facilitating the exchanges and discussions between peers during the inquiry and communicating the findings. Of course, data analysis does not always involve visualization: in quite a few scientific fields, especially in the natural sciences, data is mostly analyzed through abstract procedures and calculations that do not involve visual mapping. Throughout this thesis, when we talk about data analysis, we will be referring to *visual* data analysis, and considering the possibilities it opens for widening the publics that can profit from these tools, across disciplines and social groups.

Towards the fields of media studies and in relation to concerns like digital citizenship, we should also consider the new opportunities opened for a better understanding of the communicational exchanges outside the major mediatic encodings, of the uses this visibility

can have to foster and equip public debate and the social production of knowledge. As Viégas and Wattenberg declare, “visualization is ready to be a mass medium” (2010). We should understand what kind of medium visualization it can be and, most of all, how it may change the notion of mass communication, as every new technology does change the settings from which it is bred. Our main setting for discussing visualization is, therefore, in between scientific practices in the social sciences, professional media and everyday use. In fact, it is in the translations between these instances that reside our main concerns.

The thesis seeks to understand if in societies with ubiquitous data, that are heavily mediated by digital technologies, visualization can collaborate to compose shared spaces for public debate. Even though visualization has seen countless uses in countless environments, for the fact that we are relating it with public debate, understood as the discursive part of the exchanges in social life, something that develops shared and negotiated maps of meaning and is part of the social disputes, we will emphasize its connections to social research, media and journalism and citizenship. Visualization is, of course, also a tool for decision making, for developing corporate strategy, for public policy etc. Also, we should not ignore the fact that in these environments there is definitely much debate. Nevertheless, as we want to emphasize the political and social aspect of debate, as in public debate, we will accordingly emphasize the power of visualization for achieving innovative descriptions of social reality, recontextualizing scattered content, and providing renewed access to information sources. Lastly, as we are going to see in the first chapter, visualization itself can be a broad subject, but we will be focusing on its role in societies where data is becoming ubiquitous, as in data and information visualization. As public debate does not choose specific environments and penetrates in a way or another all social practices and the same begins to become true for data and information visualization, we do hope that some of our results may be useful for corporate decision-making, for example, but we will not approach these subjects directly.

In order to better situate our inquiry, it is first necessary to develop a working definition of visualization in the first chapter, and to discuss some of its frontiers and gray areas, according to some of the main disciplines that deal with it. This is not a straightforward task, considering that, along with the penetration of visualization tools in many spheres of our societies, there are many fields of knowledge and practice involved in its use and development. Also for the fact that our approach, for its interdisciplinarity, engages many of them. We realize that the tradition of the research in HCI (Human-Computer Interaction) does place much value on the amount of information that can be condensed and communicated

through visualization, almost as if the effectivity of visualization was mostly derived from its efficiency, from loading more and more information through our eyes at each time, and faster. Even though we understand that this is an important feature of visualization, especially in these times of data deluge, we believe this synthesis may mask a wider cultural and socio-technical context and thus narrow our understanding of the theme. For example, we identify a tendency to face data as something raw that has, by itself, relevant information while visualization would mostly reveal it. In practice, what we see is that data is never raw, it is produced inside specific methods and systems, and may carry many biases. Also, that visualization and data have a relationship of mutual constitution, which means that new and different informations may be produced from data while visualizing, and that the discrete aspect of data is paired with discrete aspects in representation systems.

Second, we should situate the fundamental importance of information access, exchange and structuring to public debate and collectivity, and outline some of the main ways in which these connections were translated into practices since the emergence of the modern democratic states. This will characterize knowledge production as a practice that is also social, that is part of collective existence, and will also point to the collective dimension of devices of mediation, among them visualization. We should also point to some issues about the circulation of messages and information in contemporary societies, in order to account for power strategies and disputes in this field. These are the goals of the second chapter.

Both aspects, the mutual constitution of visualization and data and their collective dimension, indicate that we must try to point out to which extent visualization already is, historically, penetrated by many traditions, and, from that base, to discuss how they are enlivened and reenacted in today's tools and methods. Therefore, this thesis has a double approach: the first is historical and based in discussing many examples of visual displays of information, connected to many devices and worldviews; the second was carried out through fieldwork for understanding practices and uses of visualization in data journalism and digital social research. We are focused on understanding more about visualization through some of its contexts, in two directions: on the one hand its roots, traditions and constitutive relations and, on the other hand the practices and challenges of visualization in the digital methods of social research – more specifically controversy mapping – and in data journalism. With that, we believe we have the opportunity to develop a better understanding of the workings of visualization, its challenges and the possibilities for using it to foster public debate.

In the third chapter, we develop this historical approach in the form of an antiquities hunt, bringing up examples and tracing relations, without aiming at recomposing an actual history of visualization. We advance that this mutual constitution of data and visualization has historically been organized by a progressive movement of discretization of languages, supports and technologies. This issue involves far too many levels of human craft and knowledge to allow for a thorough appreciation, but will serve as a guideline to establish some aspects of the fundamental relation between visualization and the emergence of data. In a wider sense, this emergence has been happening long before actual digital technology, and the history of visualization makes it visible, evident. This chapter is organized in five main parts: first, in a large introduction, we seek to describe this intertwined emergence of the contemporary ideas of information, visualization and of the digital; then, in the three next parts we go deeper in out antiquing and search for this movement of discretization in the representations of time and space and in the composition of context. Last, we discuss some of the main structures we found along the way: lists, tables and grids, their interchangeability and their role in the movement of discretization and its complement, the movement of reassemblage. We identify these movements of discretization and reassemblage to be part of the history of visualization and of the contemporary technologies of information, but also to be fundamental for contemporary methods of visualization.

In the fourth chapter, we address our field work in data journalism and digital methods of social research, particularly controversy mapping, towards understanding some of the main contemporary practices. In the media environment, we can perceive a tension between what would become known as the traditional model of centered communication as in broadcast media, and the possibilities now widely available through digital and networked media. Since the beginning and popularization of the internet, we have been witnessing a redistribution of content production, editing and publishing tools that also redistributes the work of media professionals and the placement of news organizations. Coupled with that redistribution, there is a growing presence of data-oriented knowledge and technology in our daily lives, as well as in the management of social life in general. Everyday, journalists face the challenge of translating, contextualising and analysing these sets of data, and one of the most powerful answers to that has been the work on infographics and data visualization. In fact, data visualization, especially for statistical and demographic data, has been increasingly present in news pieces throughout the history of journalism, so much that recently we came to see news stories composed mainly by graphs. In the emerging field of data journalism, visualization

becomes key also inside the newsroom, while producing the news pieces, as tools for analysis and for finding or building stories. For publishing these stories, visualizations have become increasingly complex, for example with online access, interactive features and real time data display.

On the other hand, scientific visualization and, more generally, scientific images, have for long been an important part of the scientific research and determinant for producing and approaching objects of study. Data visualization, while already traditional for analysing data and publishing results, comes to be especially integrated in research inquiries with the development of digital methods that draw from the data produced in digital networked media, as in digital methods of social research. By deploying links and other relations between data entities into visual structures, it opens the path to a better understanding of social networks and flows, drawing at the same time from social theory and analysis and from computational methods and models. The particular method we observed, called controversy mapping, is largely based on actor-network theory and seeks to return the maps (visualizations) produced to the actors of the controversy, as tools for public debate, to equip citizenship. So there is this alignment of social research, public debate and citizenship that is indeed very interesting for advancing the subject of this thesis.

So we turned our attention to practices of visualization in both data journalism and controversy mapping, areas of media and research where we believe the issues of knowledge production and public debate through visualization can be better identified. For data journalism, we were able to visit some of the main newsrooms in Brazil, and conducted interviews with more than twenty journalists, designers and editors from news companies and institutions inside and outside the country. For the digital methods of social research, we were able to visit the médialab, a center of research of Sciences Po Paris, and the origin of the method of controversy mapping, where we focused this part of our inquiry. We followed and participated with european project Emaps, and were able to observe up-close the challenges and possibilities of the use of visualization in the method. We also conducted many interviews with the project participants, from different universities and research institutions in Europe.

Chapter four describes the main characteristics of the use of visualizations in both fields, journalism and social sciences, details the methods used for fieldwork and the initial issues found. In journalism, we outline major changes in the use of visualizations once data analysis starts being incorporated to general practices, and relate these to new possibilities for relating

mediatic agenda to public agenda. In social sciences, there is also a similar shift, but mostly inside the digital methods of research, whose main issues and challenges we seek to describe.

In order to highlight the importance of the performative and collective aspect of visualizations as controversial devices whose many transformations may make the richness of debate visible, we draw from the work observed in the Emaps project to discuss the many movements recorded in visualizations, from data exploration towards presentation of results. We advance that we should understand this as a process of many transformations that associate data and visualization, rather than two stable and separated levels of display, one for exploration in analysis and other for communication in presentation. In order to explore these transformations, we develop two pairs of concepts: first, the reification of issues and reenaction of insights, that relate more directly to the transformations between visualizations towards debate; and, second, the camera and the album, that relate to the transformations in the actual use and access of those visualizations during the same process.

The fifth chapter is dedicated to outlining the complementary effects and strategies of *reification of issues* and *reenaction of insights*, also developing a discussion about the quest for simplification, present in the traditional uses of visualization in journalism as well as in social research. We understand that, although simplification may occur during the process, it should not be taken as a goal in itself, but as a by-product of all the transformations and refinements. Instead, we propose the idea of *shallowness*, to describe this effect of clear outlining of discursive objects and their contexts that comes from the work of reifying issues and reenacting insights. The sixth chapter is dedicated to explore the approaches of *camera* and *album*, which are two different but not mutually excluding aspects of the organization or structuring of visualizations for being accessed at each moment. We mix examples of visualizations published in media outlets with examples from controversy mapping projects. The camera approach is more evident in interactive visualizations, where the presentation is transformed along exploration of a digital reserve, and different views are outlined. The album performs the complementary movement, of collecting several kinds of visualizations into a set and sewing them together into a narrative. This brings up a discussion about interactivity, that we also address by positing that there are many levels of interactivity and many different needs for interactive resources. We understand that these should reflect the path that was outlined during the transformations in the exploratory stages, which is the fundamental and underlying data narrative.

We conclude by reinforcing the crucial role of visualization for public debate in societies that are heavily mediated by digital technologies and also for the constitution of this contemporary landscape itself. We seek to deepen the interpretation of the informations collected in our fieldwork, identifying, inside the configurations of dispute that are part of the public debate, a specific field of dispute, not exactly of the origin or property of information, but of the place of mediation of the many flows of information. This dispute sprouts from this critical approach towards the current structures of information dissemination, and does not involve specifically the creation of new and independent media outlets, but of artifacts that may be able to map out and integrate the available flows and sources of data towards new syntheses. While media outlets and information systems constantly propose structures and specific contextualizations to describe events and organize data flows, researchers and hackers extract data and make maps for new interpretations, creating new mediations. These processes revert and restructure the available flows of communication, because the data that operationally describes all the information flows is repurposed to generate new mediatic syntheses towards the interpretation of social reality, especially through visualization. To close our thesis and account for this newfound plasticity of informational flows through visualization, we propose the concept of *reverse mediation*.

2. FIRST CHAPTER: THE PROBLEM WITH VISUALIZATION

The term visualization is quite tough to frame, because it is at once broad and specific. It is broad in non-expert circles or circumstances: in the english language, the verb "visualize" may mean to imagine or remember by forming a picture in your mind, in order to organize what you know about something. When it ceases to be a verb and turns into a noun, visualization, it is linked with devices: as Ricci (personal communication, July 18th, 2014) points out, even a photographic camera can be thought of as a visualization device, because it gives you access to a specific way of seeing something. Photographs, in that sense, would be visualizations or views of that something which has been portrayed. So, curiously enough, the discussion about visualization is transferred from the realm of memory and imagination to include a network of objects and more devices that collaborate in the praxis of visual cognition. Visualization will be part of technical mediation.

On the other hand, the term is quite specific in expert fields: computer sciences, information science, a host of biomedical sciences, cognitive studies, statistics, communications and journalism, art, design and computation, to name a few, all have designations for visualization that intend to be as strict as possible, but are divergent in many aspects. The problem worsens because there are many word combinations and related terms: the frontiers between information visualization, scientific visualization, data visualization, infographics, graphs, maps, charts, visual narratives, etc. may be hard or impossible to establish. Practitioners or specialists in many different fields of research and practice will have specific terms to refer to the visualizations they use, according to their traditions and the way in which they treat their concerns. For example, many of the designers and reporters that we interviewed, who had experience in newsrooms and with data-driven news, use infographics as an umbrella term, even though the term visualization is growing in acceptance. On the other hand, scholars and designers working with controversy mapping tend to use maps as an umbrella term, even though they use quite a wide variety of visualization methods, much wider than the ones from traditional cartography.

Manovich (2010), defines visualization as a "remapping of other codes into a visual code", which brings up another aspect of technical mediation, besides the coupling with devices: the constant movements of translation. We find a correspondent discussion in Cairo (2012), who speaks from the point of view of a designer and visual journalist with large experience in

newsrooms, and interacts with these traditions, where infographics are the most widely used term to refer to graphical displays of data and information. He understands that many authors consider infographics and information visualization to be two very distinct and separated things, where the first would serve only the purpose of presentation while the second would be focused on exploration and analysis. He proposes a different approach, based on the idea that both visualization and infographics have a common nature, that of associating presentation and exploration, just in different levels. So he aligns infographics and information visualization as the two ends of a straight line (see figure 1.1), and associates the first with presentation and the last with exploration, a reorganization of terms that allows for much flexibility in the use of both infographics and visualizations, as well as experimenting with many intermediate forms. Even though we believe this to be a strong proposition, we should note that the term infographic, as it is used in journalistic circles, is also quite broad and includes presentations that are not exactly spatialized displays of data, like the floor plan of a building or a diagram that explains a food chain in a specific ecosystem. Which is probably why Tufte (1983), when discussing graphics, stresses the fact that he is addressing *quantitative* displays of information, to afterwards write another book on *visual explanations* (1997), that are not necessarily based on data. In practice, both quantitative displays and visual explanations fall into the field of infographics. As we are going to discuss ahead, information visualization, on the other hand, is considered, at least in the field of human-computer interaction (HCI), to be based on data.

Figure 1.1: The continuum between information visualization and infographics according to Cairo (2012).

Ricci (personal communication, June 18th, 2014), in an approach that is similar to Cairo's, proposes a continuum between data visualization and information visualization, also understanding that there are no clear frontiers between them. So towards one direction you would be closer to the data, and would have a focus on the exploration and the search for patterns through complex mappings, with data visualization. Towards the other direction you would have more aggregated presentations, displaying more streamlined objects and narratives and a focus on discussing assumptions, as in information visualization (see figure

1.2). Therefore, in comparison to HCI, he draws a narrower interpretation of the term information visualization, in the sense of also being based on data, but interpreted data. Direct visualizations of raw data would fall closer to the category of data visualization.

Figure 1.2: The continuum between data visualization and information visualization, according to Ricci (personal communication, June 18th, 2014).

We consider Ricci's approach to be especially useful to our discussion, for two main reasons. First, because it reinforces the general idea that information is aggregated and interpreted data, and allows us to organize some of the other terms used in many fields do designate visualization and its related formats or devices. For example, we come to realize that most quantitative displays that journalists call infographics, by associating many different contents, would fall closer to information visualization, while what they call graphs or charts would fall closer to data visualization. Second, because it also encompasses the idea that data visualization may progressively become closer to information visualization along the work of analysis, interpretation and narration, a subject we will explore in other chapters. This brings us back to Manovich's definition, that emphasises the translation of other codes info visual codes, but with a very important twist: the translation does not happen only in the first visual mapping, it continues throughout the analysis.

Of course, there are kinds of visualization that do not fall within this continuum, like most of what is called scientific visualization. Also, this schema cannot be used to trace the differences between methods of visualization like geographic maps and bar graphs, for example, because what distinguishes them is not some distance or proximity in relation to base data. But it does account for our main concerns for the discussion about visualization and public debate, which are closely related to discussing the visual interpretations and contexts for debate that we progressively build over data, through visualization. The complementary relation between the two kinds of visualization – or, we may say, two continuous aspects of visualization – makes it more reasonable to address both of them in this thesis. We will be

focusing on the work over visual displays that are based on data, be it in more direct mappings, be it in more interpreted and aggregated presentations.

2.1. HOW DIFFERENT FIELDS APPROACH VISUALIZATION

All the different takes on visualization and its related terms are, if anything, indications of the broad use of visualizations nowadays. As visualization serves many disciplines and many professional fields, sensible methods and tools for visualizing data do not cease to emerge and evolve, especially considering the growing presence of data in the description and management of every aspect of our lives. We gathered three main strands of conceptualization of visualization to which we will refer during our work: the strand of communications and design, that connects with semiotics and aesthetics and is concerned with the presentation of graphs and the messages they may convey; the strand of STS that connects to the history of sciences and some of the many methodological discussions in scientific fields that use visualization, especially where visualization allows for the use of data in social analysis, like in the digital humanities; and finally the strand of HCI, connected to computer sciences and cognitive studies, that center their concerns around amplifying cognition and visually exploring large quantities of data through interactive tools. In what follows, we will briefly introduce the main points each strand brings to our discussion, while bringing more density to the outline we are proposing for our subject.

2.1.1. Communications and design

The practices of visualization in journalistic and editorial content create an interesting space of convergence for the disciplines of communication studies and design. Designers' procedures and concerns interact with those from editors and content producers, putting visualization in a very interesting place, for its relation with different associated content, for the concern with combined narratives and with the developments in mediatic and public agendas. Traditionally, by this perspective, visualization is mostly described and approached as a form of visual explanation and visual narrative, being its most central objectives to reach visual clarity and to reveal insight. The terms used in these fields to refer to visualization may range from graphics, to graphs, to infographics, data displays and, more recently, information visualization.

After some outstanding pioneers, like the Scottish engineer William Playfair (2005), who is praised for inventing classic formats like the bar and pie graphics still in the end of the eighteenth century, and the French geographer Jacques Bertin (2011), who is known for structuring many concepts for statistical graphics for the first time in the 1960's, Edward Tufte would be the standard reference on the subject, with his work on the standards for designing proper graphical display of data (see especially TUFTE, 1983).

According to Spence & Wainer (in PLAYFAIR, 2005), Playfair started creating and perfecting some of the most traditional methods for the graphic display of data in the end of the eighteenth century, in a time when writings on economics usually displayed data in tables. Even though large archives of data had been available over a century before Playfair, it took many decades for graphics to be adopted in scientific papers. The main concerns at the time were connected with possible distortions and misinterpretations caused by visual perception. Nevertheless, outside specific scientific circles, some praised the fact that graphics could ease comprehension, like a universal graphical system that could cross language barriers. An example to this perception was the reaction of king Louis XVI, who, upon receiving Playfair's *Commercial and Political Atlas*, was highly pleased for understanding their content at once (p.1). This impression of the graphics displaying a clear and universal language can also be found in Tarde (1883), almost a century later, when he advocates that graphics could be like a clear and untainted retina for accessing reality.

In Bertin (2011), this universal aspect of graphics and visualization is still central, but will be the fruit of a careful labor of creating a graphic system to suitably display all the values and categories present in data. He understands the development of data graphs should be guided first of all by a careful consideration of the structure in the dataset, and how variables and values relate to one another. Afterwards it would be possible to test different variations and formats, in order to find the one that most clearly displays the issue at hand. He highlights the progressive and logical decisions one needs to go through while building data displays, and also the importance of graphics for information processing and storage.

Tufte, on the other hand, places much emphasis in the idea that graphs should reveal data, in the sense that, without a visual display, data and the relevant information it may contain will not be accessible. He is concerned especially with graphics as published content. He writes from the perspective of someone who has witnessed the modernization and popularization of the printing industry and the emergence of desktop publishing, and is deeply concerned with

the lack of criteria for the use of graphic resources that often followed in media outlets and even in scientific publications. Tufte advocates clarity, precision and efficacy, and points to the need of clearing the graphs of what he calls *chartjunk*. Chartjunk are unnecessary and confusing decorations or visual metaphors included just for the sake of visual impact or entertainment. In opposition to this problem, he coined the concept of *data-ink*, which refers to the "non-erasable core of the graphic, the non-redundant ink arranged in response to variation in the numbers presented." (1993, p.93) So one of the central purposes of Tufte's principles will be to maximize as much as possible the share of data-ink in relation elements devoted to things other than data, the non-data-ink. Even part of the data-ink itself may be removed (and its corresponding data as well), if it is considered redundant to the issue at hand. Tufte also creates a measure for *data density*, based on the proportion between data points displayed in the graphic's matrix (total entries) and its total area, in order to treat the issue of a graphic's relevance (see figures 1.3 to 1.5).

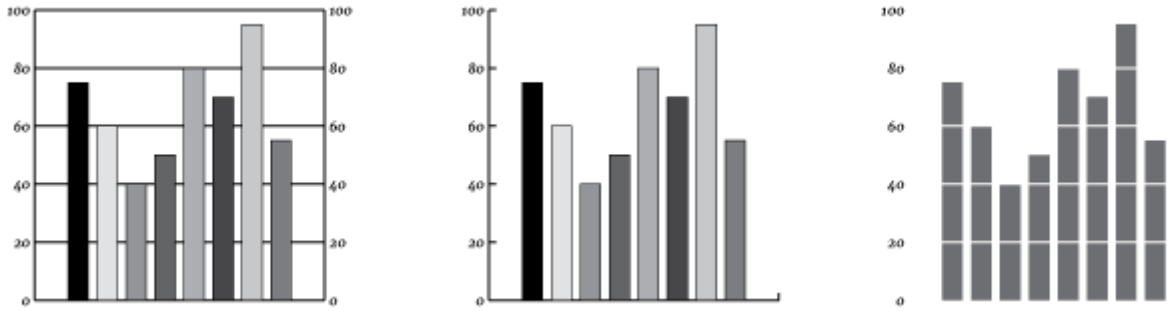


Figure 1.3: From left to right, an example of progressive improvements on the data density of a bar graph by removing non-data-ink elements. Based on TUFTE: 2001, p.126-127.

There is also another set of concerns, which is the relation between graphs and other kinds of content and the possibilities for developing visual narratives through data and information. Tufte understands that combinations with other kinds of elements, like textual comments, labels and highlights, collaborates for making data complexity accessible. Like Bertin, he believes the fundamental choice of how to display quantitative data will be guided by the characteristics of the dataset, like the amount of labels and the volume and ordering of data. But, at the same time, as he is concerned with publishing, he highlights that there are different basic devices for displaying data: sentences, text-tables, tables, demi-graphics and graphics. For reduced amounts of values or depending on the kind of comparisons you wish to draw, graphics might not be the best choice. Graphics will in general be the most powerful choice as datasets become more complex, but even when they are chosen as the main solution for the

display, the final published result will usually combine other devices, like sentences and tables, in order to attract and introduce the reader to the richness of data. For Tufte, a certain narrative quality of graphics is important to their clarity, and is reached mainly by the care with these associated devices.

Tufte also notes that there is an artificial separation between images and text in modern publishing, that he credits mostly to the separation of professional specialties. He believes this hinders the proper integration of content, be it in the page structure or in the content producing itself, where graphics are published as illustrations and texts become unnecessarily heavy with redundant data content. Kanno (2013), on the other hand, demonstrates how historically the printing technologies collaborated to such separation: typesetting, for example, served mostly for text, while initially it was necessary to print images in separate pages, using different techniques. The technical possibility of integrating image and text, two parts of a same content, was only partially achieved with off-set printing, after a long process of technical improvements that was simultaneous with the structuring of journalism as a professional field, so a strong heritage of separation between images and text is still very much alive in practices of page layout, reporting and editing. Kanno proposes the term "visual journalism", highlighting the integration of various visual resources within the text to better tell the news story.

In recent references, and also in Tufte's more recent works, different concerns take central stage, like the problem of larger and larger datasets with many variables. That more complex data displays, based on large and multivariate datasets, may fall into the concerns of communications and design is no mere accident, considering the growing presence of data for the description and management of today's connected societies, followed by the interest of wider publics, that feel they are affected by issues that are better approached through data analysis. To better address this, there are deeper investigations on the possibilities of exploratory and interactive visualizations that would enable the display of many details and different views while keeping a sense of context for displaying the data as an integrated whole. Before, this kind of tool would be mostly used in visual data mining and other activities related primarily to the analysis of data and not for displaying relevant issues for wider publics. So the field of communications and design starts dealing with different aspects of visualization, with a tendency to combine visual exploration and visual explanation. For example, Cairo (2012), who is specialized in infographics and has a large experience in newsrooms, talks about a partnership between exploration and presentation, pointing to this

double use of visualization in the media, something that wasn't very much highlighted in old school references in the field.

This shift in approach signals to an approximation with some concerns of statistics and computer science, fields that traditionally have worked to advance robust tools for the exploration of complex datasets, but rely on tried and tested formats and lack much concern for the visual quality of data displays and how effectively they may be communicating. Fry (2004), for example, advocates a deeper methodological integration between visual design, data mining techniques and software-based information visualization, in order to deal with the limitations in each area. So there is a growing and mutual influence between practices that work to aid understanding and to improve communication, and others that focus on complex analysis and can handle the large amounts of data present in today's networks and information systems and organize the infrastructure for that. A sign of this growing exchange is also the fact that in professional practice in communications and design the terms information visualization or data visualization, that are traditional inside computer science and human-computer interaction, are growing in acceptance.

Finally, one last relevant concern: Tufte (1983) and Cairo (2012), for example, advocate the refining of the graphics in many cycles of progressive revision and editing. As visualizations gain complexity and new metrics for evaluating user behavior are available for online publications, these cycles can encompass refining the graphics according to the interaction of a wider public of users. In that sense, we can profit from the specific field of design and their research on methods, even if we do not have many references addressed directly at the problem of visualization. While the tradition of design have focused historically on delivering form that follows function, recent theory and methods will be concerned with addressing the many situations of use and, moving ahead, with the different functions (and meanings) that may arise from use (CARDOSO, 2012). The designed object gets to be considered something that evolves in time, going through many unexpected appropriations. In terms of design methods, this is translated into iterative cycles that encompass an attention to the uses before design, and the uses after design: the appropriations should be included into design practice in order to feed new cycles of improvements and transformations, as in participatory design and in the idea of design after design (DISALVO, 2012). We believe these efforts tend to approach the designed object as a collective, social object, and not as something that is projected for the cognitive and functional needs of an individual, for individual reception and use.

Something similar occurs in the field of communication studies, where, apart from the theory directly related to visualization in newsmaking, we find the specific subject of visualization to be somehow underdeveloped. Nevertheless, we believe we could definitely profit from the classic concern of the field with the processes of mediation and with the reception by readers. Recently, and especially with the arising of some issues more evident in digital media, communication studies have been offering relevant keys to discuss visualizations as collective mediation devices and as interfaces for broader social processes. Indeed, we believe approaching visualizations as media devices and including it as such inside the field's concerns is one of the main contributions we intend to bring with this thesis.

2.1.2. STS and the history of science

The field of STS approaches visual documents in general as evidences of scientific processes and methods. They are also seen as tools for the communication between colleagues, that make collective research practice possible, as well as demonstrations and reinforcements of scientific ethos, as in the work of Daston & Galison (2010). On the other hand, they are analysed as traces of the many stages of the work being done, of the development of discussions that generate scientific knowledge, where the transformations and translations that happen along the way are in focus.

Daston & Galison (2010) discuss the many historical developments of the quest for objectivity in the natural sciences since the eighteenth century, using the scientific atlases as empirical object. They highlight the collective aspect of the atlases, and the fact that they aimed at "selecting and constituting 'working objects'"(p.19), being tools for "collective empiricism"(p.26) that train the eye to recognize significant traits and frame what is acceptable as a scientific object at each field of the natural sciences. For the authors, scientific atlases display the evolution of different approaches to scientific practice and different criteria for perfecting fundamental epistemic virtues, among which objectivity. And these criteria draw the boundaries in which scientific debates take place.

Eighteenth century atlases, carefully crafted by reputed naturalists and their chosen artists, collaborated to refine ideal types of nature, that is, seeked ideal representations that would clear out many variations between specimens. It was believed that the variations masked the underlying nature of the phenomenon: the authors call this perspective "truth-to-nature". By the end of the nineteenth century, mostly with the first experimentations with photography, it

becomes clear that the variation in specimens is part of nature, part of phenomena. After the evidence delivered by mechanical record, all the efforts of refining archetypical types start to be seen as far too human and, to put it briefly, subjective. So atlas makers become obsessed with intervening the least possible in the images produced mechanically, so that nature could be portrayed without the distortions caused by human idealistic projections, interpretations and expectations. The machines, rather than direct observation, become instruments for scientific discoveries, in some sort of blind sight. They call it "mechanical objectivity": the objects portrayed become characteristic, examples of certain issues, in order to draw, by comparison, a perimeter of the variations of what was to be considered normal and what was deviant. But as the images multiplied, so did the criteria of organization, the passion for detail, hindering the clarity and transparency that they were aiming at.

As Daston & Galison advance, this perspective evolved into two upcoming views: one that seeked to abolish images themselves, called "structural objectivity", and another that valued visual interpretation, called "trained judgment". The first one, structural objectivity, refers to groups of scientists that were concerned with the limits and variability of human perception, and also of language. They believed that reality and experience could never be fully represented and transmitted from mind to mind, because: "anything that was picturable, subject to the laws of association, and above all private was *ipso facto* 'psychological' and could not be modified by the adjective 'objective'" (p.270). The challenge here was to "transcend the privacy and individuality of representations and intuitions" (p.273). So, for them, the quest for objectivity involved searching for things that could be identical for everyone: they believed only pure relations could be communicated and were, therefore, the only objective experience possible. Simple structures for collective convenience were to be the goal of every scientific work. The true essence of things was out of reach. The closer to images this scientific ethos created were abstract schemas that aimed at communicating these relations.

On the other hand, trained judgment brought to light the limitations of using specific images as "boundary posts for the normal" (p.309). The fact was that atlases like the ones from mechanical objectivity originally proposed that the task of representing should stay on the image itself, but towards the twentieth century, this responsibility was increasingly transferred to the readers. It was in the hands of the reader to compare, relate and arrive at new syntheses, so judgment followed from there. Scientists that work inside the ethos of trained judgment will advocate and assume the need of a discipline for judgment and interpretation that would

add to the mechanically produced image. They point to the "necessity of seeing scientifically through an interpretive eye" (p.311), and to accept the role of intuition in scientific hypotheses and scientific work. This new training and discipline for judgment would produce an "interpreted image" (p.311), refined by trained judgment, and end up building a new set of criteria for atlas making.

Abstract schemas, describing experience and reality as a set of relations, recognizing and analyzing patterns... Unfortunately for us, Daston & Galison pass right beside the subject of information visualization, but do not dwell in it. Nevertheless, this massive theory does address many relevant issues for treating the subject of visualization: it establishes how several kinds of sight were involved historically in building scientific knowledge and how this is a collective and therefore communicational problem. More specifically, it also hints at the emergence and evolution of visual comparisons and relations as well as visual patterns as important assets for scientific research and communication.

While Daston & Galison focus mostly on the finished atlases or images as objects of analysis, Lynch (1985) presents us with a discussion about visual documents in the sciences that focuses specifically on the processes of selection and mathematization images go through so they can present and take part in the construction of scientific facts. For him, by the progressive use of selective perception, images are annotated and schematized in a sort of distilling process until they can clearly display and even demonstrate the object of interest and its relevant characteristics. He makes the point of declaring that it is necessary to separate the idea of observation of its individual and cognitive base, and associate it with what he calls "rétine exteriorisé"[1] (p.110), that is, the shared material where images circulate, that helps building working objects to be dealt with. For Lynch, in the traditions of pragmatic, phenomenological and interpretative sociology, perception is many times compared to a filter that, from a chaotic world, operates choices, simplifies and puts in order, according to projects and interests of each individual. He believes we should regard the idea of selection as something that is connected to coordinated practices of collectives, and not only to the sole individual. Therefore the idea of externalized retina: it is the visual criteria that is externalized as social practice in scientific images.

[1] "Externalized retina" (our translation)

This is not a process of invention: these images are, in fact, a construction that emerges from the interaction of several actors and procedures. He describes two main themes: the simplification, that addresses the way in which scientific methods simplify and schematize their objects of study, and the mathematization, or the movement of attributing mathematical order to natural objects. To discuss simplification, he brings up pairs of photographs and their correspondent schematic drawings: one is the result of the initial object being turned into image by an automatic photographic device, the other is the result of a process of reduction and filtering of the first one to its relevant graphic elements, like contours, frontiers or surfaces. So there are sequential transformations that aim mainly at: filtering noise and irrelevant detail; standardization to better fit visual conventions and better communicate; amelioration, to improve the clarity of the image; and définition, for clearly distinguishing different elements, which entails classification to a certain extent. As for mathematization, it entails a geometrization of forms that would better clarify measures and reinforce the approximative linearity, uniformity and regularity ascribed to the objects along the study (see figures 1.6 and 1.7).

Latour & Biezunski (2005) talk about the hardening of scientific facts as a social undertaking, that is, many methods and processes by which facts are outlined, proven and reified, many times, to an extent in which they become irrefutable. Much earlier, in an article titled “Les vues de l’esprit” (1985), Latour had already put forward the idea of the *mobiles immuables*, or immutable mobiles, which would be in the center of scientific activity: normalized and stabilized objects that could be recombined and transported to different contexts. Later, in the book Pandora’s hope (1999) he also beautifully describes a series of transformations between the complex (or raw) object of study and progressively compatible and standardized elements, in devices, inscriptions and images, where each keeps a verifiable reference towards the previous, allowing for the retracing and verification of procedures.

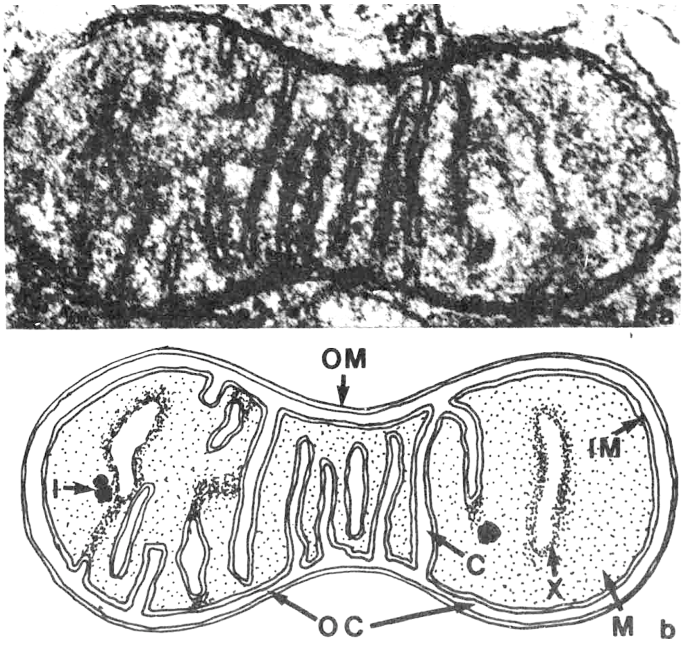


Figure 1.4: *Ribosome model*. Simplification in visual scientific documents. Source: Stephens, Grover; North, Barbara. Biology. New York, John Wiley and Sons, 1974, p.87. *Apud* Lynch, 1985. p. 112.

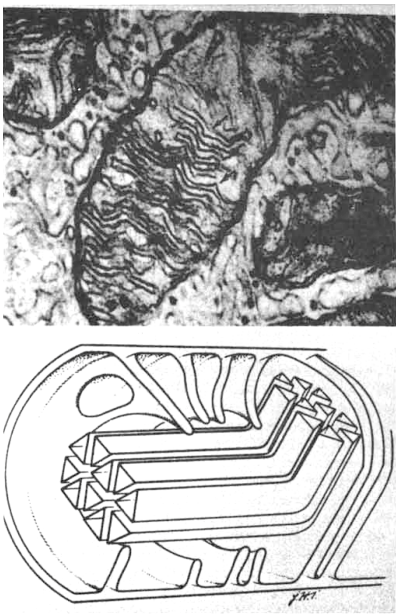


Figure 1.5: *Ridges*. Mathematization in visual scientific documents. Source: *idem*, p.327. *Apud* Lynch, 1985, p.115

So, as we have seen, most of STS literature, when discussing the role of images for the construction of scientific knowledge or even of different epistemologies (like in Daston & Galison), deals mostly with what we could call scientific visualization. There is, like in the trained judgment described by Daston & Galison or in the externalized retina of Lynch, a coupling with automatized and systematized methods of documentation and layers of annotations and refining that consolidate elements. Latour does include a wider set of devices and methods alongside the specific activities of image-making in the transformations he follows, but the spatial character as a reference that can be trailed back to the original messy natural object across each step of transformation is still an important issue.

As we intend to further develop in this thesis, the spatial character of information visualization is ascribed and built in a different way: since in the beginning of the visualization process what you have is data, there is no image of nature (neither of external visual references) to be refined into other images and image sequences. Spatiality is ascribed according to methodological issues, hypotheses and relations between quantities and/or categories in base data. And a certain spatial character may be traced all along the

transformations or may change completely between images: as we discussed earlier, testing different forms of spatialization is key to the process of visual data analysis. And this is not necessarily a difference that goes back to the separation between natural and social sciences: natural sciences do use information visualization and, on the other hand, social sciences might use maps that target physical distribution of social phenomena. With digital measuring and sampling devices, translating enormous amounts of records into data and information visualizations has become commonplace. To an extreme, part of what is registered in the Big Hadron Collider at CERN, for example, is analyzed as data, through visualization: either because particles are so small they cannot be actually photographed (or otherwise imagetically registered), only measured; or because they are tracked by such a huge amount of indirect traces, that it is necessary to assemble them into visualizations in order to recognize patterns.

This leads us to a very delicate discussion, which is the differentiation between scientific visualization and information visualization. Munzner (2008), for example, understands that the difference lies in the ontological origin of the spatial dimension: while with information visualization the spatial dimension is chosen (or at least configured or constructed), for scientific visualization the spatial dimension is given. This can be related to the process of "imposition", as described by Bertin: in information visualization there is always the fundamental stage of deciding how data will be mapped, of imposing spatial dimensionality that was not there to begin with. On the other hand, Bederson & Shneiderman (2003), researchers from the field of Human Computer Interaction (HCI), provide us with a more detailed and, we believe, more accurate differentiation. They understand that researchers of scientific visualization are concerned primarily with three-dimensional objects, with their volumes and surfaces, fluxes and formations, and seek to answer questions about inside or outside, above and below, left or right. On the other hand, researchers of information visualization are focused in representing abstract phenomena to which there might not be a correspondent physical reality to be captured as an image: variations on the values of stocks, monitoring of production and distribution flows, variations in glycemic levels, social relationships etc. Even if both sources are coming from the physical world, with information visualization the emphasis is in finding relations between variables and patterns in the whole, while with scientific visualization reproduces objects and phenomena in images in order to understand physical structures and processes by means of annotating and measuring.

And this also has effects on the kind of data and variables used:

> "scientific visualization users are primarily interested in continuous variables such as density, temperature, or pressure, whereas information visualization users deal with continuous as well as categorical variables, such as gender, race, home ownership, date of birth, state name and number of bedrooms." (BEDERSON & SHNEIDERMAN: 2003, p.ix)

These differences in the variables used also points to the proximity between information visualization and the social sciences, that will need devices to deal with phenomena that are not immediately visible.

For Lynch (1985), scientific visualization builds the physiognomy of scientific objects, as the movements of simplification and mathematization transform samples in facts. Likewise, we could say that information visualization, while working from complex data mappings towards new syntheses, builds visibility to different objects, and turns samples into measured phenomena. One starts with data, the other starts with an image, but both will face many challenges to ensure the reliability of their starting points and of the following transformations. So, instead of drawing the line between scientific and information visualizations by evaluating which has a more direct or indicial link to a reference, it seems more productive to realize that the conditions for the emergence of spatiality are different at each case.

Of course, like much of the STS literature we are using elaborates, all the annotations, tracing, filtering and improving that scientific images go through until they turn into stabilized working objects will eventually make them closer to diagrams based on measurements than to pictures of the raw and messy natural object; closer to pictures of what is known about something rather than pictures of a specific something.

After all, should we scribble yet another continuum, this time between scientific visualization and infographics? We probably could: at one side, we would have photographs or their variants, indicial impressions of a certain specimen. Towards the other, we would have, progressively, annotations on their surfaces, schematic drawings that would clear much noise and define relevant limits and surfaces, and progressive regularization and geometrization of forms, until we got to a point where the scientific image would not be discernible from – picture that! – an infographic. We might as well come back to figure 1.7. What matters most to us is that in both continuums – the one between data visualization and information visualization and the other, between scientific visualization and infographics – the road seems to be gained with progressive interpretation and outlining of relevant objects or elements. The

focus for these transformations seems to rest on in pattern recognition, either for pulling apart different textures or borders, or identifying sets of occurrences, in order to separate information from noise and to generate the interpreted image.

So there are two main concerns in STS regarding the scientific images or inscriptions that are central to our approach. First for associating images to the collective aspect of scientific developments, for providing science with shared stabilized objects that set the stage for exchange, communication and collective work and, second, for providing access points for understanding the development of those discussions through observing the visible transformations between images. While in STS these concerns are more clearly approached through discussions about scientific visualization, we feel they are very relevant for understanding information visualization and its role in creating visual working objects. We understand that the process of deriving scientific knowledge from visualizing data entails specific concerns and will aid collective work, stabilize shared objects and go through many transformations in very specific ways that merit much attention, especially considering the role of data in today's connected societies and, on the other hand, the penetration of scientific controversies and evidence in public debate. As Lynch, for example, problematizes the relation of scientific visualizations with the pictured object, we should problematize the relation of information visualization with data, and the transformations elaborated over it.

2.1.3. HCI and computer sciences

HCI scholars, on their turn, use only the term "information visualization" to refer to the spatialized presentations of data that are in the core of their work as interfaces for the interactions between humans and computers. Information visualization, in fact, is considered to be a subfield of HCI (BEDERSON & SHNEIDERMAN, 2003), so the field offers a quite strong and organized literature on the subject. Card and other relevant scholars coined a widely adopted concept of information visualization: "the use of computer-supported, interactive, visual representations of abstract data to amplify cognition" (CARD *et alii*, 1999, p.7). This concept makes clear that there is a deep concern for the technical features a visualization may dispose of and also for the subject of amplifying cognition. Tying information visualization to those technical features (of being necessarily interactive and computer-supported) will seem quite restrictive for the ones coming from the theories of communications and design and from journalistic practice, but they also signal to the focus of

the field on complex mappings and exploratory tools for visual data mining, where they deliver the most useful theory to our work.

We should note that there are authors in related fields who do not observe this strict link between visualization and interactive features, and sometimes not even with computer-supported technologies. Börner & Polley (2014), for example, use the most varied examples of visualizations in their work, from interactive, to digital static, to print, to hand-drawn. Even if they focus on advanced data mining and computer tools, and even if Börner identifies herself as a researcher and professor in Information and Computer Sciences, they take into account the insights provided by many different resources for the task of making sense of data, and have a broader view on what visualization is. We will discuss and refer to HCI's contributions for the theme of information visualization, but will take on a broader approach, like the one proposed by Börner & Polley.

In HCI, this general goal of making sense of data is translated in the concern with amplifying cognition. In fact, for HCI, the theme of amplified cognition is in the backbone of the relation between humans and computers, where there is a complementarity between human visual-cognitive system and computer graphics. Ware (2003), for example, highlights the importance of thinking tools for cognition: the capacities of the mind for reasoning by itself are overrated, because thinking almost never goes on solely inside people's heads. We think with many objects, we think with the world, so cognition is externalized with instruments, and amplifying it involves creating better thinking tools. Information visualization is, for HCI, an interface between the two systems of computation and human cognition, and the relation is two-fold: if there is an amplification of cognition, there is, on the other hand, an amplification of the computing possibilities through human cognition. A strong example of this is visual data mining, where humans are said to *lend* cognitive power to computers, so data can be organized (KEIM, 2002).

According to Ware (2003, p.3-4), the main advantages of visualization are: allowing the understanding of large amounts of data; giving a general view of many records, therefore allowing the perception of patterns or emergent properties; making problems with data (like biases and vices in methods or artifacts) evident; facilitating understanding of both large-scale and small-scale features of data; and facilitating hypothesis formation. Apart from that, the author does highlight that visualization fundamentally takes advantage from the visual cognitive system, that is capable of processing several data points simultaneously, registering

visual stimuli in parallel. Card *et alii* (1999) summarize this last aspect by praising the bandwidth and the processing capabilities of the visual-cognitive system.

For Card *et alii* (1999), visualization is one of the procedures by which one can develop a process of knowledge crystallization, that is motivated by the search for a better understanding of a problem, with the aim of feeding decision-making. Knowledge crystallization involves different stages: first, there is information gathering and the construction of a representational framework to make sense out of what has been collected. Using a framework will always involve the schematization or abstraction of information, as well as the omission of informations that are not relevant to the problem and do not fit into the chosen schema. This means that, on the one hand, the problem is formulated while knowledge is crystallized, and also that in this process there is always some loss, some cutting off of information on account of the normalization of data. So there should always be some consideration for the trade-offs of the available options. From there, with information structured in the chosen schema, one can package the patterns, that is, highlight the patterns discovered and instrument the discussion or propose a solution for the problem.

This evaluation for fitting the data into a schema is clearly a process of reducing an initial only broadly defined problem into a matter of pros and cons, a logical issue, it is driving the problem away from a much uncertain and undefined terrain. Card *et alii* highlight the selective character of this process: the task or problem at hand will orient the omission of complexity levels in the data that are not relevant in a given context. They also point to an interesting theme, that the selective omission of information is at work in every biological information processing system, as part of mechanisms geared towards the efficiency of actions.

It is also interesting to notice how this process of knowledge crystallization describes visualization not only as something that develops over time and is paired not only with decision-making processes, through interaction, but also as something that is related with the physiology of visual cognition itself, at least as it is described in the field. This points to the integration of the human cognitive system and the computer, that is so dear to the research in HCI. Bederson & Shneiderman (2003), for example, do bring up the importance of the *flow* in information visualization. Flow is a concept developed by Csikszentmihalyi (*apud* BEDERSON & SHNEIDERMAN, 2003) to address the sense of an optimal and fluid experience in one's exchange with the world. For information visualization, this is translated

into a few objectives and recommendations, so the computer interface can allow the users to concentrate on the tasks at hand, becoming an extension of the user's body and offering users a sense of control and confidence while he/she makes progress toward their goals (p.XV).

Ware (2003), translates the search for this integration between information visualization and user into seeking a sort of alignment between the human cognitive system and the computer, so he developed a set of recommendations for information visualization based on the research on the physiology of human cognition. According to him, visual and cognitive processing happens in three stages. To put it briefly, first, low-complexity data like shape, color, orientation, movement and texture is collected from the visual field at the same time by the many sensors in the eye, that work in parallel. This is stored in a temporary iconic memory. In a second stage, there is a slower process of devising different regions in the whole that has been perceived, building a map of the distribution of those traits and its patterns. On the third stage, objects are identified and stored in the short-term memory. Long-term memory interacts with the more superficial levels of iconic memory and short-term memory all the time during this process: according to pragmatic problems the individual is faced with, the brain translates tasks in visual parameters and performs visual queries throughout the system. So the process combines bottom-up procedures in the low levels of visual cognition with top-down categories in the long-term memory.

Therefore, there is this pairing or partial mirroring of cycles in both parts, of cognition and visualization, where both systems activate the whole chain through interaction, guided by practical needs or interests. But, more than that, we identify that they are both (even human cognition) described in terms of data extraction and treatment, while the social environment seems to be left outside of the many interactive cycles. As we are aiming at discussing the uses of visualization for public debate, it can become quite limiting to talk about cognition mostly as a physiologic, and therefore individual, process. From our point of view, it is not enough to consider that data itself comes from social environment and the insights developed through the joint cycles of cognition and visualization go back to it. Even if Card *et alii* (1999) work with the concept of external cognition, in a certain sense cognition is brought back inside the interiority of a closed circuit with a specific device. The schemes of the authors we referred to encompass a one-on-one relationship with the machine, but we might just need to address the amplification of cognition in other terms: it will have to include debate and collective reasoning also as amplifications. And, of course, this cannot be done

without some consideration to the bumps, irregularities, translations and compositions that come into play when discussing technical mediation inside social practices and public debate.

2.2. THE WORKINGS OF DATA AND INFORMATION VISUALIZATION

French geographer Jacques Bertin (2011) advocates from the beginning that thought can only be expressed through a system of shared signs, and that graphics should work just like that: as coherent systems of signs to communication and thought. Therefore, the title of his classic book, The semiology of graphics. He proposes many terms that aid evaluating and discussing datasets from the point of view of graphics-making, a task he calls the analysis of information. From the structural understanding of data that these terms collaborate to, there is a process of *imposition* of data into the two-dimensional plane, that we could call a first spatialization of data. Before everything, this involves the choice of a graphic structure that should follow the structure of the dataset, more specifically the relations between variables (that he calls components). Of course, since the geometrical plane will initially display two dimensions, one at the vertical axis, another at the horizontal axis, two main variables of the dataset must be chosen for this initial imposition, even if more are available. The other dimensions may be portrayed by different features, such as color or size of graphic marks, or even by projecting a third dimension on the plane, or displaying variations in small multiples. So, in these two variables, if there are correspondences to be traced between all the values in one variable and all the values in the other variable, then the graphic should be a diagram. For example: a bar graphic that displays variations on the prices of a certain commodity, in which a price corresponds to a specific date. If there are correspondences between all the values in the same variable, it will be a network, like in networks of social interactions. Finally, if the values of one variable are to be distributed on the plane according to their geographic location, then we have a geographic map.

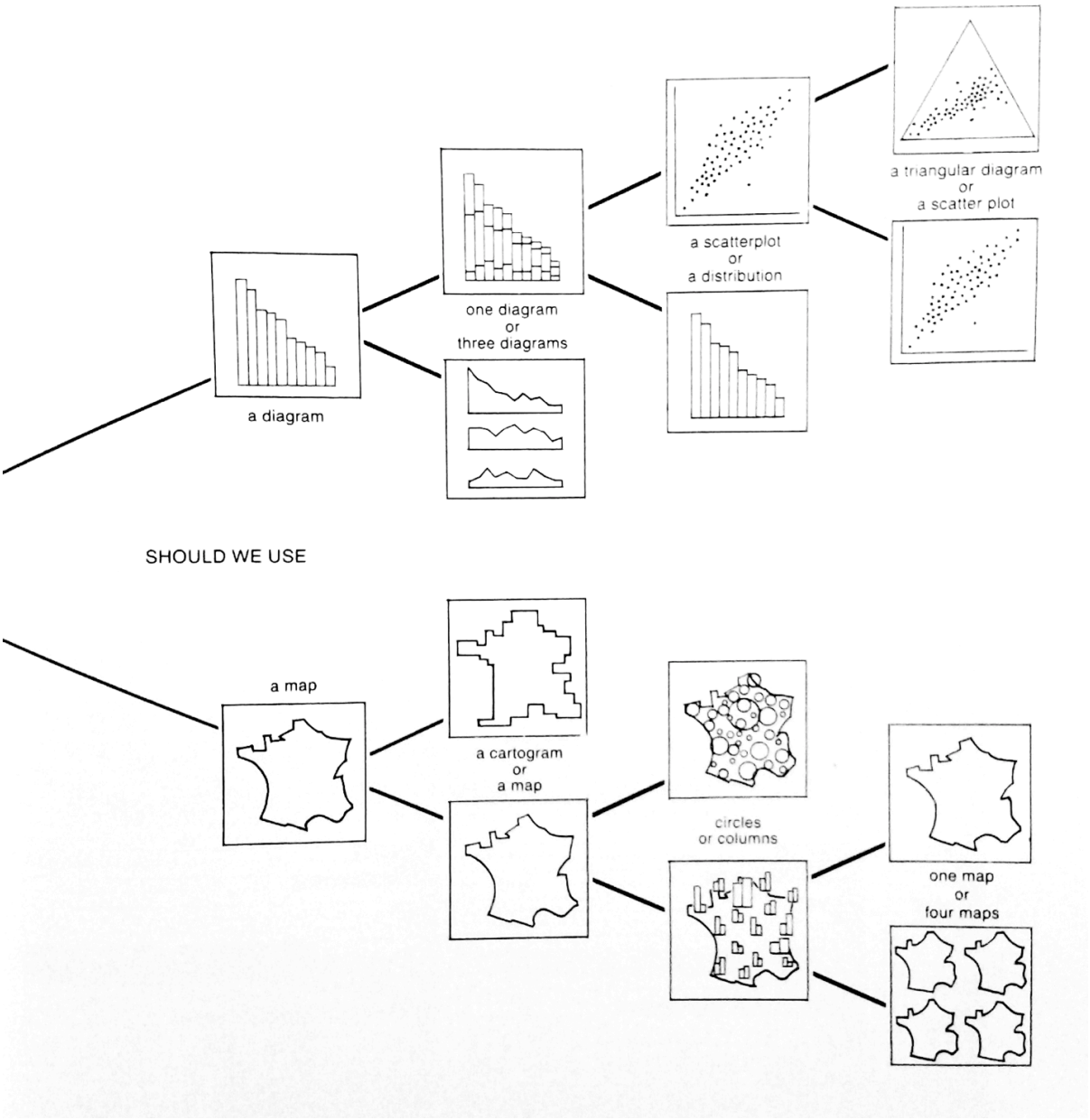


Figure 1.6: A tree of possible choices for mapping the same data table. Source: BERTIN:2011, p.101.

It is important to notice that these criteria, however straightforward and data-adherent they may seem, also involve considerations that are not strictly connected to data structure, but to the aspects one may choose to highlight on data, according to the problem at hand, and also to concerns over the situations of use and clarity for publishing. This can be perceived in an example he develops to discuss other more detailed choices of graphic structure, after this first level of analysis and decision. Even in a dataset that contains geographical data, structures other than maps can and sometimes should be used (see figure 1.9). There are many different

graphic solutions to address different data and varying concerns in analysis and emphases in display.

Börner & Polley (2014) also point to the relation between graphic structures and the internal relations between variables and value in the dataset, but highlight the different questions that can be asked at each configuration of data and graphic structure. For them, there are five main types of analysis, that are related to questions that can be asked: Statistical Analysis/Profiling, related to traditional demographics and so on, drawing profiles and categories from a group; Temporal Analysis (when), that deals with the evolution of entities in time; Geospatial Analysis (where), when the focus is on displacements, locations or distributions in physical space; Topical Analysis (what), focusing on themes, word analysis etc; and, finally, Network Analysis (with whom), that focuses on relations and interactions between entities. There are also three levels, depending on the number and coverage of records being analysed: from micro/individual, to meso/local, to macro/global. This categorization that is centered on kinds of questions to be asked and the size of the universe to be approached highlights the focus on the insights that the activities of building visualizations provide during and after their undertaking.

They propose a workflow for the design of visualizations (see figure 1.10), in which at start you would have the stakeholders (providers of the data, readers or concerned publics), then (as in Bertin) you would go through a first superficial analysis of data, to determine its general structure and coverage. At this stage, data will be *read* and *analyzed*, and it might have to be cleaned and preprocessed. Afterwards, you get into the stages dedicated more specifically to visualizing, which are choosing a visualization scheme; overlaying data, that is, inserting (or imposing, for Bertin) the data into the chosen structure; and, finally, visually encode the data, that is, working on standards to represent the other dimensions that may work well together. Once this visual system is defined, the visualization can be deployed – that means to be published or distributed – and presented to stakeholders for validation, which will initiate new cycles. They define this as a needs-driven workflow design, in the sense of building visualization according to specific needs or tasks, that will guide from data selection to parsing, to graphic structures and the interaction with them. No wonder it is organized as a loop that integrates the stakeholders.

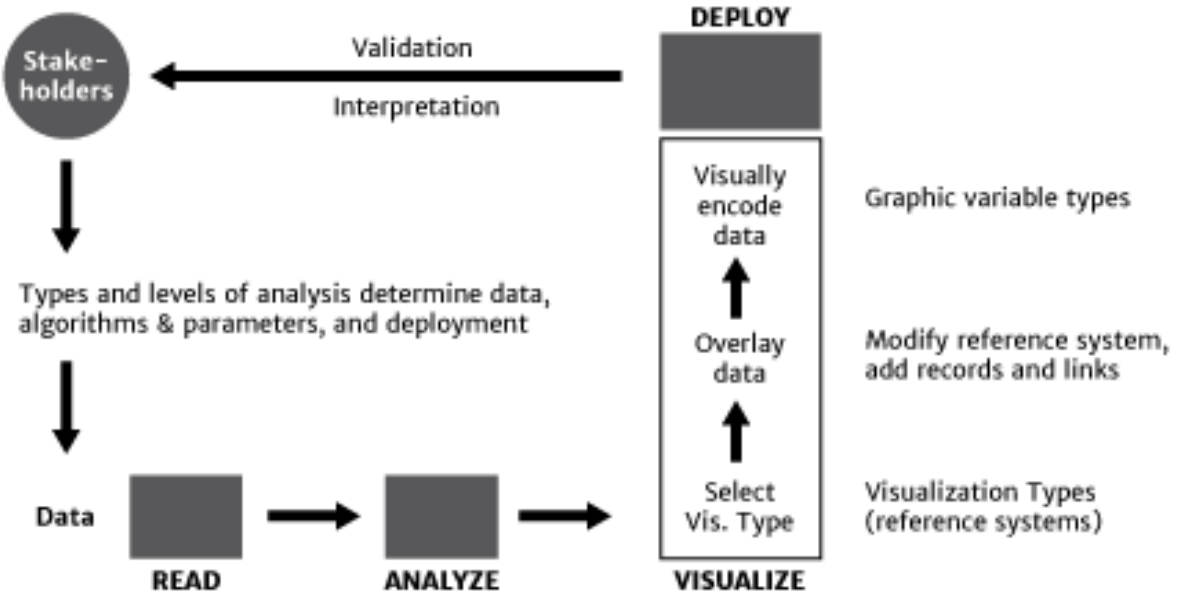


Figure 1.7: Needs-driven workflow design. Redrawn from Börner & Polley, 2013, p.18.

One other interesting aspect of the work done by Börner & Polley (2014), is that they do see visualization as a kind of framework for visual exploration and insight, but do not link exploration necessarily to interactive features. They understand that even static or print visualizations can be relevant for visual data analysis. According to them, there are many kinds of visualizations, related to their structure and level of complexity, and each one will be best for addressing a specific kind of question.

Card *et alii* (1999), on their turn, define a reference model for visualization (see figure 1.11) that gives special attention to the transformations imposed on data and the feedback loops performed during the process of building and accessing visualizations. It is relevant to us for offering a view on the specific processes for interactive visualizations. From the start we have raw data, that may have the most varied presentations and are in this sense idiosyncratic: registries that may be hard to combine or simply are not structured according to specific needs. To work around these variations and arrive at a normalization that allows for the visualization of uniform levels of comparison, data is inserted into tables. This involves a transformation of data to reveal objects and relations that may be described in mathematical terms, and the combination with structural data, or metadata. This is an intermediate stage, between raw data and visualization, when data becomes standardized, sortable and workable, and at the same time assume already a first spatial characteristic that is fundamental for performing basic comparisons.

Then there is the mapping of data into graphic structures, that afterwards incorporate spatial coordinates, visual highlights and other graphic properties: all of these create a code for visually presenting the the dimensions, quantities and categories present in data tables. These graphic structures are like visual information models, that are going to be displayed according to interaction in updated views, that is, transitory organizations of the structure that are produced according to the user's exploration. The passage between different views is achieved by view transformations, that are the result of transitions demanded by user interaction. As a matter of fact, the visualizations HCI is interested in are processes of visualizing from data, to tables, to graphic structures and to views, they exist in time.

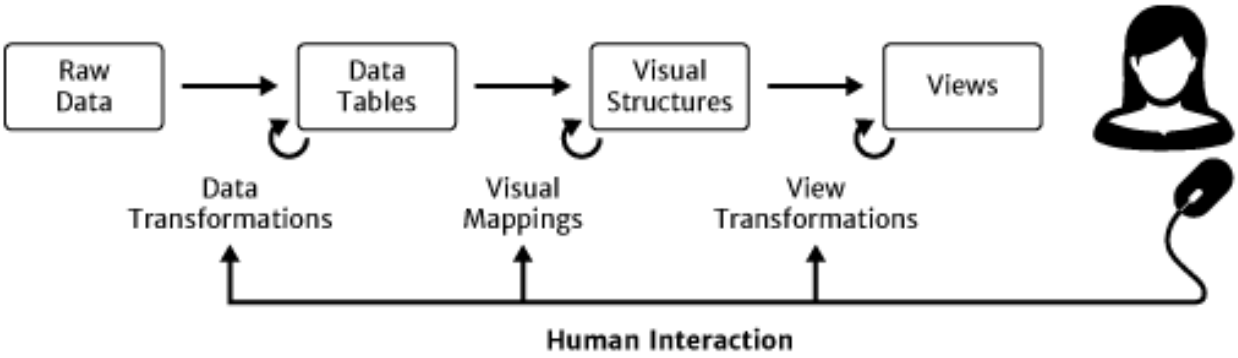


Figure 1.8: Information Visualization Model (redrawn from CARD *et alii*: 2013, p.17)

Card *et alii* have also classified some kinds of transformations and described some levels of feedback loops. After the data transformations involved in transforming raw data into data tables, which may include discarding values or making approximations for missing ones, there are the transformations directly associated with imposing data into the visual structures, and, finally, there are the view transformations, which are something exclusive to interactive visualizations. There are basically three of them: location probes, in which extra data related to a specific data-point or area is displayed, adding to the data present in the initial landscape; viewpoint controls, where the visual structure is magnified or displayed by another point of view or another part of it; and, finally, distortion, which distorts the larger structure in order to show detail while keeping sight of the overview in a single continuous space. The view transformations, conducted during human interaction, account for the feedback loops we see in the reference model (see figure 1.11). We believe it is because of all these necessary transformations in data that they choose to use the term information visualization in the place of data visualization: it implies a formatting and interpretive work so that visualizations may exist, a process that extracts and maps relevant data, to relate it towards information.

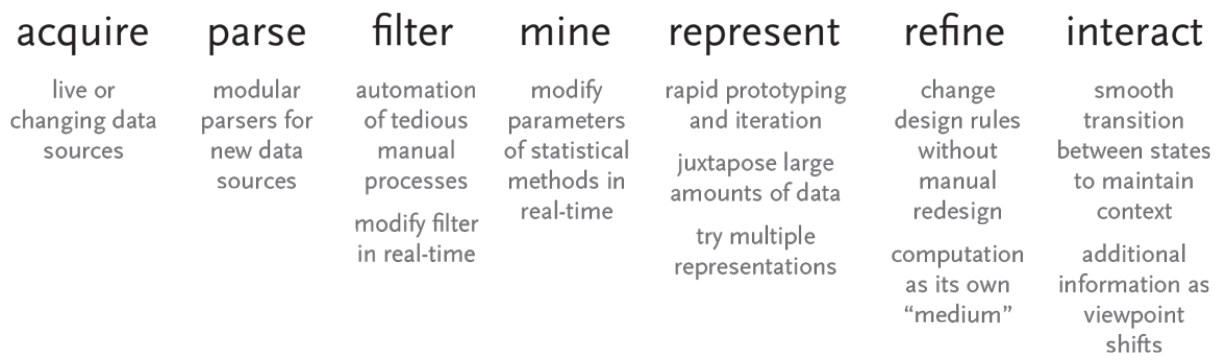


Figure 1.9: Stages for producing a visualization. Source: FRY: 2004, p.31.

Fry (2004) points to the fact that information visualization encompasses tasks that mobilize many different kinds of specialists, and proposes a method that might better integrate different capacities, that he calls *computational information design*. As he argues, the different stages of information visualization (see figure 1.12) are usually best conducted by different kinds of professionals: acquire and parsing are inside the field of computer science; filtering and mining, in mathematics, statistics and data mining; representing and refining in graphic design and, finally, planning and deploying the interactive tools is a task for information visualization specialists and HCI. Every specialty has some limitation towards information visualization: visual design usually does not have the tools for handling large amounts of data; data mining techniques are meant for dealing with large amounts of data but do not address the means for interacting with it; software-based information visualization usually builds tools or features for interacting with large datasets, but is disconnected from the visual design principles that would allow for better understanding and operation (p.11). He also believes that, as every professional specialty deals with a part of the process isolatedly, "each step of the process diminishes aspects of the initial question under consideration" (p.14), so much density and interest are lost in the final result.

Like the model from the HCI scholars, Fry's model considers the cycles of transformation in user's interaction, that reorganize the chain, so production and access are overlayed. But Fry, for bringing to the center stage the exchanges between professionals with different expertises, takes one step further: for every stage, he proposes a few methodological improvements that add to the visualizations' flexibility in terms of new materials and criteria that may emerge along the process. At the acquire stage, the system should be able to deal with live or changing data sources; at the parse stage, modular parsers should be built so new data sources can be included; filters should automate processes and allow for real-time modifications; and

so on and so forth. So, instead of many different systems, one for each visualization and dataset, Fry brings up the modular possibilities of programming as a way of working towards building blocks that can be integrated, recombined and reconfigured for specific and changing needs. We believe this is an interesting way of technically incorporating ongoing debates and the collective aspect of research in the core and workings of visualizations. Of course, this can be related to many practices and methods that do not cease to gain acceptance in programming and software development community, like widely adopted standards, data-exchanging interfaces like APIs, and open-source software. Between the extremes of situation-specific visualizations and the utopic (and perhaps badly framed) goal of producing universal or all-purpose visualization systems, we find Fry's perspective to be more productive, at least considering interactive computer-supported visualizations.

2.3. ASSUMPTIONS TO BE QUESTIONED

As we have seen, most of the traditional literature that can be more directly linked to data and information visualization is quite in accord about a few assumptions. First of all, they praise the fact that visualization will condense large amounts of data into a single image. Visualization is, in general, valued for the *quantitative* advantages that are ascribed to it: mostly for the fact that it can aggregate and represent at once enormous amounts of data and for the fact that it takes advantage of the capacity of the visual-cognitive system for also transmitting these large amounts of information at the same time, in parallel. From our point of view, these advantages tend to be taken as straightforward values in themselves, probably for the many and relevant concerns raised by the data deluge we have been living in since the popularization of digital technologies and networks. We believe we should allow ourselves to question the fundamental character of these advantages, so we can contextualize them while exploring the constructive relation between information flow and visualization.

Another advantage or quality that is ascribed to visualization is that it reveals information that would otherwise be hidden (or we could say invisible) in data. This, again, is built atop some very reasonable assumptions: checking and comparing the huge amounts of records that can be fed into and produced by today's information networks is impossible without some method for aggregated and contextualized access. Comparisons between different individual records would be impossible in simpler structures like tables and lists, so some relevant information would not be uncovered. But, on the other hand, we feel that this phrasing of things tends to

take the relationship between data and visual mappings for granted, as something natural, and also tends to assume that data has in itself, a certain amount of information that only needs to be brought to light, to the visible world. This second assumption is probably the most dangerous, because it could lead to the idea that data is some sort of hidden or underlying code for everything and visualization comes as a new tool that would help us uncover it. It could also lead us to disregard the fact that data is a form among many to describe things for specific uses, and especially that visualizations are after all interpretations that may lead to many other different interpretations in analysis, and also carry many rhetoric strategies that build over data. We should keep in mind that this polyphonic, interpretative aspect is in the roots of producing information and especially of building knowledge.

We believe the emergence of data as such a powerful and useful description of our reality goes back to a long process of *discretization* in the techniques and technologies for recording, storing and exchanging knowledge and information. As we intend to demonstrate, this can be traced in workings of devices and in the surfaces and structures of documents and images, and points to a joint development of data and its spatial presentations, among which there is visualization. So the contexts we build from data through visualization should be seen as something a bit less normalized and straightforward; and the relationship between data and visualization should also be seen as rich in tensions and productive interaction that may lead to unexpected results.

This leads us to another major aspect of the traditional literature about visualization, which is its collective aspect, that is, in our opinion, underdeveloped. As we have seen, transformations are always highlighted when one discusses making and accessing interactive visualizations, but they are not contextualized inside collective practices and methods where visualization is inserted: transformations are not seen as part of the debate. Nevertheless, even as researchers, for example, work with static visualizations, these serve the purpose of communication together with analysis, in the sense that people gather around them to discuss and exchange ideas. The different versions of visualizations, especially of the static ones, are not only by-products of sequential improvements in the design process: they are the trail debate leaves behind, and testify for its collaborative development.

Connected to this concern for the collective nature of visualizations, there is also the problem of how this body of theory deals with the issue of insight in visualization: it is often seen as an individual enterprise, connected to the efficient functioning of cognition. Again, this sprouts

from some very relevant and undisputable discoveries of cognitive sciences about the functioning of human visual perception, but, as we would like to discuss below, the most efficient cognitive and perceptual system may not be the most insightful. For the rest of this chapter, we will take a closer look at this specific issue.

2.3.1 Sight, interaction and insight

Indeed, sight is a very complex sense and a very powerful dimension of human cognition. According to Bederson & Shneiderman (2003), "There is simply more bandwidth and processing power for input through the human eyes than through any other sensory modality." (p.ix) It offers the possibility of non-linear apprehending through images, whereas the other senses would only dispose of linear or sequential modes. So in a surprising equivalence to computation devices, HCI researches would say that vision offers the possibility of parallel processing while the other senses would offer only serial processing.

Nevertheless, we believe it is important to understand that, if this is true, it is so partly because vision has gone through a long process of specialization and coupling with different visual practices. Crary (1990) describes this long process where former functions of the eye were progressively altered by a detachment between visual and bodily experience. Especially from the first half of the nineteenth century on, there is a significant growth in "practices in which visual images no longer have any reference to the position of an observer in a 'real', optically perceived world" (p.2). Vision gets abstracted and conversely abstracts what is seen. It is so common to find praises for the parallel processing of the visual-cognitive system, that we forget that other senses also offer this possibility. It is key for hearing, for example, otherwise we would lose much of our spatial orientation. For example, according to Chion (1985), the introduction of Dolby sound systems in cinema rooms was crucial for our sense of immersion in the image: while vision has been largely abstracted from bodily presence, the other senses continue to be closely connected to our bodies and *ground* our experience. In cinema rooms, sounds that come all around the audience re-locate the body into a tridimensional space, collaborating for an immersive experience with the film. So the same technologies that profit from all these advantages ascribed to vision collaborated to form ways of seeing that that are useful and productive for them and, at the same time, were themselves developed along centuries for this vision that was emerging.

On the other hand, the fact that parallel processing opens the possibility of making comparisons very efficiently is key for the uses of visualization, but it is also central for our interaction with the world in a general way. According to Ware (2003), once we know our surroundings, we have developed cognitive schemas, or mental images derived from experience, that summarize the most useful aspects and the points of interaction with the environment, to which we cling to the most. Comparing only superficial perceptions to those general schemas in order to get around lowers the pressure and the effort of being always attentive, always focused, open and actively processing detailed information. Ware explains that the process of visual perception and cognition in three stages that we described earlier, as well as the process of rechecking the whole chain with the images stored in long-term memory and new indices from vision that tunes up the whole system, can all be understood as a process of *attention* (p.22). We are always performing this as we go through our lives, and attention can have a superficial, fluid aspect to it or it can be active, selective.

We understand that visual interfaces in general encourage active and selective attention, not because this is necessarily the best way to explore them, but because they are designed with hierarchical graphic systems to differentiate elements and a priori visual relations, at least in a superficial level. Also because, especially in the case of interactive visualizations, we can say that their functioning is like a chain of actions and reactions: commands from the user and returns from the software that updates the views. Therefore, when Ware seeks neurosensorial evidences for the development of visualizations that may be quickly learnt and understood, he privileges of course a selective and focused use of vision, that receives clues each time clearer and stratified, that aim at clearer communication and easier operation for an acceleration of interactive cycles.

The problem is that, if the predefined schemas of the sensorimotor system reigned entirely, there would be no space for the uncertainty that leads to discovery. It is necessary that there is some break in the flow of interaction for the discoveries to take place, otherwise visualizations would only be communicational tools in the most restricted sense, of transmitting some pre-existing information, and would never do for investigation or analysis. If something new is discovered, it is because it differs, it strikes out and breaks the flow. This is related to the idea of breakdown, as proposed by Varela (1989): when some inconsistency is perceived, this leads to an inventive break in the existing perceptual structures, it is the key to invention. We would like to point to the fact that many times the success of visualizations is gauged by how much they integrate with cognitive habits and collaborate for their

efficiency: if routine cognitive procedures aim at saving effort and resources to be efficient, likewise, the geometry used in visualizations flattens a land to be accurately measured and mapped, it takes off unexpected bumps. If insight is in the deviations, we could run the risk of using visualization to obscure information and insight instead of revealing it.

According to Deleuze (1990), indeed we normally perceive only clichés: very efficiently, we process everything we perceive to create *sensorimotor images* that are partial, that record whatever is useful at some point according to our interests. This simplifies decisions and lets us focus on possible actions. Of course, this agrees with the researches referred to by Ware (2003) and points again to the issue of selection and loss to guarantee the efficiency of cognitive and perceptual systems, present some way or another in all biological systems. From a certain point of view, images also serve this purpose: they are visual systems that give away some codification so that we do not always perceive everything, because indeed that would be unbearable. Every image ends up by falling into a cliche. Even when we perceive something for the first time, never seen before, the hyper population of memorized sensorimotor schemas is such that they will work as filters, so that many details will recede to the background.

Deleuze posits that it is necessary that the sensorimotor structure breaks from the inside so that we can have a more direct access to what we see. It is right in the moment when these schemas break or are not functional anymore that another image emerges, a whole image without any metaphor that might cover it, the direct image of the thing itself, in its radical character. This breaking is the key to the experience of the *image-time*, a concept created by Deleuze to address many aspects of the aesthetics of cinema. For him, in cinema history there is a passage towards the experience of the image-time from the neo-realism on. In traditional realism, the cinema image has a strong link to the sensorimotor system, in such a way that the viewer, identifying himself or herself with the characters of the drama, extends perception into action, following the chain of events in the movie. The sensorimotor image occurs in a *well qualified* (p.14) environment, that is, clearly established, and works with the game of action and reaction: action that is going to reveal more about the environment or that will react to it, either aiming at adaptation to it or modification of it.

Interactive technologies are generally thought of in terms of efficacy for getting things done and problem-solving, which favors the continuation of the image into motor action. In fact, we believe a common effect of interactivity is this emphasis in the action-reaction cycles, that

are in general described as a three-step sequence: overview, displacement (pan) and view of details (see, for example, KEIN, 2002). In general these cycles reinforce this motor aspect of perception that removes us from the object and what is seen and, at one extreme, may cover gaps, irregularities, perturbations or simply empty areas that are necessary for opening space for problematization and investigation. Deleuze encourages us to look at the the cycles of access of visualizations through a different light when he describes the possibilities for overcoming the sensorimotor system in the perceptive cycles themselves. These are described as cycles that always return to the landscape that has been formed, producing new interpretations, amplifying and reassessing at each time the field of references that is associated to it, from where mental images are mobilized. So at each cycle, the access to details, instead of reducing the landscape by framing, may amplify the information field that is being mobilized, and covers the mental image that was formed in the previous cycle with another full image of the whole. We will try to see the interaction with visualizations as the access to a small virtual reserve that is reconnected in cycles of progressively larger coverage.

Deleuze takes reference in Bergson and the differences he traces between motor (or automatic) recognition and attentive recognition: the first kind of recognition operates by prolonging (or projecting) cognitive automatisms like visual schemas into the image, and therefore habit and everyday uses would reign: perception is prolonged into usual movements and the movement prolongs perception to take from it useful, practical effects. This ends up by removing us from the initial perception, and creates a horizontal movement of association of images and accumulation of schematic and motor data. The second kind, on the other hand, the attentive recognition, there would be an interruption of this prolonging of habits, and the attention would be freed to return to the object and discover further details, restart over and over from scratch, and the movements would have a different nature, would be more subtle. This separation shows how the passage from one form of recognition to another – associated to the rooted schemas on the one side and their overcoming on the other side – is in fact quite floating, it is a variation from one to another.

At this point it is important to note that Deleuze does not propose that this exit from the sensorimotor schemes and their continuation in useful action represent an exit from action as a whole. He understands the action happens, but takes on a more mental aspect. Bringing once more this discussion closer to the context of media, and, therefore, with the more encompassing discussion about the relation between subject and technique, we would like to bring up the difference, traced by Kirsh & Maglio (1995), between pragmatic action and

epistemic action. While performing an experiment with players of video-game Tetris, the authors managed to determine that not all the commands that were entered by users aimed at better positioning the the pieces in the most adequate position to fit with the bottom ones. The game demanded that many actions should be performed in a minimum amount of time, and therefore a maximum economy and efficiency of actions should be expected for good results. Nevertheless, many users, especially the most experienced, at times moved and rotated the pieces just to think, regardless of the best position for each situation. Even though the game offered a limited quantity of fitting positions and those very quickly get to be memorized by players, it was like they seeked to visualize the contours of the piece and the possibilities for fitting it directly on screen instead of mentally projecting these possibilities. According to Kirsh & Maglio, this kind of action can be characterized as epistemic, because it manipulates the world in order to think more efficiently, to form a visible field of possibilities and diminish the demands over memory and cognition. We understand that visualization should work as an environment where the emphasis should be on epistemic action, that have no direct pragmatic goals, and at each cycle will amplify the field of information that is associated and mobilized.

3. SECOND CHAPTER: PUBLIC DEBATE AND VISIBILITY

As we have been discussing, data and information visualization are used in many different fields of knowledge and practice. Even as many references will point that their main advantages are dealing with and transmitting huge amounts of data by their coupling with the visual cognitive system, we want to bring into focus other aspects that may be more influential for public debate and collectivity. In this chapter we intend to trace some relations between data and the management and representation of social life. Also, considering that the access of information is key for participation in democratic societies, we also intend to discuss the role and importance of making the social visible through data and information visualization.

One thing is certain: we have never produced such huge amounts records. In the last couple of decades, most of the theoretical and critical works that have been dealing with this issue start with the same *adagio*: impressive statistics that show that the amount of information being produced is growing exponentially, and that just in the last decade – or even year, or month – we have produced more information than in the whole history before the twentieth century (see, for example, WURMAN: 2000). We are also reminded that all this deluge of data completely surpasses the cognitive capacities of a human being, that is, a person cannot pretend to read, observe and understand all the content that flows from different sources, not even by far.

For the present discussion, first we could point out that once one cannot process a comprehensive amount of information about the society he/she lives in, this will obviously interfere negatively on his/her participation in collective concerns and, to an extent, make this person more vulnerable to the many competing forces that might affect his life. For that, many counter-measures have been proposed: Wurman (2000), for example, presents us with the growing field of information architecture, for taking on the task of structuring and editing information, even before digital networks, and embed published materials with visual navigation aids. Others will point to the growing importance of filters and relevance sorting that might reduce the amount of information to deal with directly, or at least prioritize the different contents, so that people can find their way in a complex and scattered information landscape. Producing these interfaces or devices to organize the access of information becomes one of the main concerns of fields like archivology, information sciences,

communications and others. Does that mean that the data deluge in which we are living now is making it harder to reach reliable and comprehensive syntheses of our experiences and of social life? At the end of the day, will the data deluge and the broader access to information that we have now only allow for either too complex or too reducing representations of the world, while keeping reality itself, as an experience, out of reach?

These questions may be quite relevant from a personal and psychological point of view, but, at the same time, we believe we should look a little deeper and try to understand upon which assumptions they are built, asking ourselves yet another set of questions. For example, what would it mean if a person could actually process all the information and data produced nowadays? Before the digital age or even before typesetting, were we ever able to do this, to process every possible information we could potentially get our hands on or reach with our senses? As we discussed earlier, the cognitive system is highly efficient, in the sense of producing mental images or cognitive syntheses that will guide us through known landscapes and serve as references for any new *stimuli* that comes our way. Things are taken for their correspondent schemas unless significant variations strike us or unanticipated needs emerge. In other words, the human cognitive system is very efficient precisely in *not* processing everything that comes its way. So why is it that, especially with digital technologies, desktop publishing and, later, the internet, we feel like we are obliged to process enormous amounts of information just in order to participate in the reality we already live in?

Imagine a person, in the seventeenth century, for example, who has always lived by the side of a forest, and must cross it to get to the nearest village. If he/she knows all the main paths that cut through the forest, can locate its main landmarks, like a stream, some bigger rocks, maybe a clearing or even some very significant trees, this person might be said to know the forest quite well. But, really, he/she does not need to have a comprehensive memory of each changing leaf or every growing tree, nor does he/she need to know the parts of the forest that are outside the already marked paths in order to reach his/her goal. This person only needs to have their roadmap memorized.

Now imagine that the forest is constantly changing its major physical features and paths, and that, in order to cross it, this person must rely on contradictory descriptions left by different people who crossed it recently, sometimes retold by other people who did not cross the forest themselves, added to traces left by animals in their non-related activities, that should be interpreted on the spot. How much information would be necessary to build a mental image of

the forest that is updated, reliable and sufficiently detailed? How does one decide which pieces of information are the most reliable and which should discarded? So this is where our metaphor is heading: our biggest anxiety is perhaps not having a reliable roadmap to all the intricacies that affect our lives. And this may grow more and more distressful as globalization and communication systems create productive and cultural links that extend far beyond our immediate reality, and that we feel that affect us just the same. Because of that, it is hard to recognize when we have built enough knowledge in order to be fair in our judgments, safe in our investments, successful in our tasks, sure about the relevance of our concerns and so on and so forth. Potentially there is no limit to information gathering and, without constant gathering, there is also no criteria to update and assess the reliability of our provisory maps.

So even though the amount of information produced nowadays poses many challenges with which a whole industry of information and communication technologies is very much concerned, we would like to highlight that this is the surface of deeper and older concerns, that is brought to the forefront once we take the role of information in today's heavily mediated societies as our point of reference. The production of data grew and will keep on growing, but this is because data and digitally-based information became structural to research, to communications, to most technologies of everyday life, and so on. Data-based forms of information organize our lives. Digital technologies work like an universal converter, not because data descriptions can get to the essence of things, but because they are useful for many ends. And this leads us to the uses of data for the management of social life and for the establishment of shared references: social reality is penetrated by data and information that are used strategically to organize its many voices. Information devices (especially digital, but not exclusively) are technical tools but also set the arenas where the negotiation of collective interests takes place. So as concerned and affected individuals we would do well to keep up, developing our own strategies for information consumption.

Of course, the belief that broad public debate is needed for fairer societies and also in its necessary connection to the broad access of information carries many assumptions that derive from ideals inherited from the Enlightenment. These went through many transformations and reinterpretations since they were developed in the foundations of modern democratic systems. On the other hand, Marxist tradition will integrate debate and information production and exchange in a wider setting of social struggle. According to Whiteford (1999), recent reinterpretations of Marx's works will place much emphasis in communication as a force that may produce resistance and highlight autonomy even as strategies of control penetrate the

tiniest of social relations. To better ground the discussion that follows, we would like to look a little deeper into some of these currents and interpretations, so we can identify some main issues regarding the dynamics of public debate, and trace some links between these and the management of social life through data, as well as communication technologies and the media in general.

3.1. REASON, INTELLECT AND PUBLIC DEBATE

While discussing the famous article by Kant, *Das ist Aufklarung*, Foucault (2009) describes the Enlightenment as some sort of exit, where people would go from a state of *minority* when the use of reason and reasoning would be restricted to a few, to a new social order where individuals are otherwise aggregated by the public use of reason and reasoning. They would leave a situation where they would be forced to delegate the use of reason to an authority that would lead them, where there was no room for the private use of reason, or for reasoning as a human activity in general. In this movement of exit from the restrictions to the public use of reason, there had to be a collective and individual choice to take on the task of daring to know.

For that to happen, first it was necessary that one could pull apart what ensues of obedience and what ensues of reason: develop an understanding that, inside a regime where obedience is required, it would be possible to make use of reason. Obey to be able to reason, and reasoning, in this sense, for the purpose of reasoning itself, as a human capability that integrates humanity. Second, it was also necessary differentiate the public and the private use of reason: according to Foucault, Kant understands that reason should be free in its public use and submissive in its private use. Here, the private use of reason corresponds to its use in the private sector, meaning in work, production, when the individual is part of a productive apparatus where tasks must be performed in a certain way, with specific ends. So the use of reason cannot be freely exercised, and should submit to these circumstances. In this sense, individuals should understand the political principle of obedience and from this understanding would obey, willingly collaborating towards the broader social harmony, without the need of the use of force or any forms of social coercion. On the other hand, the public use of reason occurs when one reasons as a member of rational humanity, and then it should be free. From this we understand that the Enlightenment was not just about guaranteeing the freedom of thought and reasoning for people in general, but of constructing the grounds that would make

its universal use coincide with its free use and it public use, in the effort of preserving private men from authoritarian and arbitrary attitudes of their sovereigns. As Foucault points out, in this text Kant understands that even the political principle of obedience must agree with universal reason, to which every authority should bow, and only then can obedience be rational.

Foucault sees in Kant's standpoint an early rehearsal of what could be considered a modern attitude, which involves: a search for a heroic present, demanding some questioning about how one relates to it; a creative work of imagination and transfiguration of this present; and last, devoting oneself to their own reinvention as an autonomous individual and historical being, reinventing and *aesthetizing* life itself. This would describe, for Foucault, the modern solution to the exit initiated by the lights: in this sense we are all historical beings, still determined to a good extent by the Enlightenment. Nevertheless, there is a remarkable difference between these solutions: while for Kant the Enlightenment was the epoch of the critique, that is, of constant evaluation about where or to which extent to use reason, Foucault proposes what he considers a positive turn towards investigating the limits of the use of reason and the possibilities for surpassing these limits towards new experiences. We see, in this turn from a negative perspective towards a positive one (in Foucault's words), a shift on the public use of reason, where it should become, in other words, investigative, led towards experimentation, not and not exactly towards parametrization.

What interests us most about these reflexions is how both of them place the use of reason in the center stage of social life, and it comes to be deeply linked with discursivity and debate. Each vision, in its own way, advances the general idea that reason is experienced discursively, and that the task of building a common social world demands opening space for discourse, debate or exchange. While for the experience of the Enlightenment reason should be the universal measure of the meanings and limits of authority and obedience, for the moderns, it would come to be the common ground for transforming and re-narrating the present towards the future. They are both talking about liberation, but different kinds of it: for the Enlightenment, it was an exit from the subjection to the sovereign, an affirmation of reason as a human trait and the foundation for building of a safer common world; while for the moderns it was a different exit, that could be described as a way of living – emancipation as the only possible destination.

This general perspective that associated reason with the constitution of public spaces through discursivity, and in a sense devised discursivity as a form of social praxis of reason, will be in the roots of the modern democratic state, as proposed by Habermas (1991) in the concept of the public sphere. According to him, the public sphere is a rational and discursive space with relative autonomy of the state, where private men can discuss and negotiate matters that affect the collectivity. The space of negotiation and exchange unravelled in the public sphere derives mostly of some transformations on the participation of the state facing the growth of private institutions in the public scene. In a monarchic state, the power is concentrated in the person of the king and his image, and so his will is absolute, not needing to resort to any authority or external reference, nor any socially shared notions. The status of this power is indifferent to the categories of public and private, as are all the feudal or pre-bourgeois power structures.

As we summarized, Kant had already proposed that, with the Enlightenment, there is a separation between public and private, while universal reason should be general reference for obedience and negotiations. Focusing on the formation of the institutions in modern democracies, Habermas describes the long process of disaggregation of centralized feudal powers until the end of the eighteenth century, with a polarization of private elements at one side and public ones at the other. That is how the institutions of public power are differentiated from the private sector, and are progressively converted in state bodies of nations, putting together public administration and relating closely to the military power. So, in the roots of the modern project of democracy there is the result of a search of the bourgeoisie for autonomy in relation to the power of the state, in the establishment of a social space ruled by reason, called public sphere. It emerges as a space that could guarantee a relative emancipation of the private affairs, including the domestic, familial and psychological spaces, also comprising of course commercial exchange and the productive sectors.

Habermas also highlights the discursive aspect of the collective use of reason, defining the public sphere also as a field of social life where public opinion is developed. The later derives of an argumentative practice and involves the exit of reason from the psychological spaces and from domestic intimacy into public space. Public opinion is, therefore, the social product of the public use of reason and should guide social practices and the negotiation of private issues that affect the collectivity. For Gomes (2005), public opinion is the will that legitimizes itself as reason (p.19). So, in this sense, the public sphere becomes some sort of publicization of reason through public opinion.

According to Gomes (2005), this concept of public sphere, as exposed by Habermas, demands that, in the modern democratic state, there has to be a certain degree of rationality, where several points of view would mutually and honestly verify one another, aiming at reasonable convictions, standpoints and consensus about the issues being discussed (p.16-17). It is a formal and polite debate between private gentlemen. So, if this notion is used for describing modern states it becomes quite naive and in practice surprisingly excluding of several parts of the collectivity that are directly affected by the issues being discussed. Naive because it assumes that it would be possible, at a large scale, that all citizens were willing to openly debate issues of interest without resorting to authoritative discourse, equally devoid of pride and vanity in relation to their own convictions, and without imposing personal interests in detriment of the collectivity. And excluding because there is a clear issue of access, once the ideal of the public sphere seems to be closely connected to an ideal of debate between like minds and well read people, which would demand some ability and even a specific discipline for debate, let alone disposing of private property and the means of production. That is why, for Gomes (2005), Habermas' concept of public sphere should be taken as some sort of prescription, that is, an indication of democracy's fundamental values and duties, rather than a description of its actual practices. So this prescription, this proposal, is going to be put in practice in different ways according to context.

It is true that the principles of the Enlightenment looked at reason as a universal and basic human trait that would inform all authority and obedience, and that this was presented in that historical context as path for liberation of humankind through conscience and willing participation in duties in society. Nevertheless, the interpretation of these same principles by recent authors such as Habermas may betray a point of view that, even if seen as prescriptive, demands a discipline for debate, fences it up in some specific arenas, and fails to consider the unavoidably polyphonic, non-uniform and even chaotic character of public debate, if we intend it to be more inclusive.

Indeed, as Rancière (1996) argues, in the greek Polis – that was a fundamental reference to democratic ideals from the Enlightenment on – the basic exigence for politics was a presupposition of an equality from the point of view of speech, that would mirror an equality of possibilities in a debate. Since those times, this equality of speech demanded a scene to be set, a common ground for speech. This equality also was, by all practical terms, impossible to be entirely fulfilled: for the ones who do not master the use of language – or argumentation and rhetorics, or knowledge of the disciplines of the management of social life or that are not

in a position of speech –, all that is left is to yel. Rancière notes that equality will become its contrary once it is transferred to a place of social and state organization, that is, institutional (p.46). Therefore, this intellectual emancipation that we have been discussing in the terms of the exit described by Kant and of the posture of the moderns according to Foucault, as it also appears in the ideal of the democratic state in Habermas' concept of public sphere, cannot be institutionalized without becoming some sort of instruction of the people, that is, organizing and managing their permanent state of minority.

So when Habermas goes on to describe a degradation of the public sphere in the second half of the twentieth century through mass media and through the penetration of private interests in public institutions, we have to point out two main problems: first, that the same public sphere can be considered, from the beginning, as a space for the management of states of minority, in which collective life is organized according to the vision of likewise minded and educated gentlemen. Second, that, in this sense, mass media and public institutions only embody means and strategies of domination specific to a certain epoch, a few among various others, that are transformed along a history of permanent dispute. In both aspects, what we see is that we cannot fully separate debate and collective interest from the mess of the streets and popular struggle without pairing it with strategies for domination. The public sphere is, in this sense, degraded from its very start.

This takes us to one more aspect of the configuration of the modern western democracies, especially from the second half of the twentieth century on, that is the staging of the public sphere. Both political institutions and the media work towards formatting political positions and developing strategies to ease public adherence to them. So in modern democracies the public is occasionally called to participate mostly by choosing between this or that position, represented by this or that party or candidate, while the only part of this institutionalized public sphere of debate to which they have access is generally staged to present these reified discourses and positions. To Gomes, from the twentieth century on, and specially with mass media, positions still have to be defended discursively, but now this is not done *inside* the public sphere, but staged *before* the public sphere (2005, p.23). So the public sphere becomes, not a medium of debate from which one might expect to see the emergence of an opinion, but a medium of circulation of established opinions to which adherence is expected, as broadly as possible, from a public that is called from time to time to act as in a referendum. A public sphere that is constituted by these artifices is nothing more than a medium of propaganda (GOMES, 2005, p.24).

We could point out many spheres where public debate is staged and these positions are organized strategically, through political campaigns, media discourse, public policy, marketing and product management, among others. This reduces the complexity of discussions in state politics, in current affairs and in themes approached by the media, but also of individual experiences when they are grouped in demographic categories, of the tastes and preferences one might cling to in identification processes. From our point of view, it is important to place public debate inside and beyond state politics and the traditional idea of the public sphere, and include in our discussion these other more distributed strategies, that aim at managing something more pervasive, to which Gomes refers as the *public scene*. This reduction that operates in the most varied levels of collective and individual experience is often necessary to allow for the management of several aspects of our society, but to an extreme it does limit the complexity of debate and with that the possibility for participation in collective matters. Gomes calls all these efforts of management of the public opinion an *engineering of consensus*. In that sense, we might be in dire need of an *engineering of controversy* that could be more inclusive.

Marxist tradition, on the other hand, centers the description of societies at their conflictual aspects, at the social struggles between productive and living labor and the efforts for its domination by capital. Whiteford (1999) describes what he calls the deaths of Marxism, by the critique of mainly two currents, the Neoliberalism and the Post-Marxism, and moves on to reaffirm the multiplicity of Marx's works and the relevance of Marxist currents for understanding current societies. For Whiteford, the neoliberal critique does not differentiate marxist theory and state socialism, and places the end of both in the disintegration of the USSR and in the assimilation of China into the globalized market. According to it , one of the main reasons for this collapse would be an inherent incapacity of socialist regimes to deal with new information technologies and post-fordist production techniques, that some neoliberal voices would argue could be traced back to some inherent flaws in marxist theory. Post-marxist critique, on the other hand, would point to what is seen as a reductionism on the part of a more traditional marxist theory: that people's identity should not be reduced to their role as workforce, that domination is carried out and identities are formed in very different ways according to specific social groups, such as women, immigrants, children etc. So classical marxist theory would be reducing people's complex interrelations and identification processes to their capacity for work and to the struggles related to waged labor, while restricting spaces of struggle to the workshops or factories, and not fully considering the

participation and the complexity of other spaces such as the ones where unwaged labor is executed, like female domestic work.

Dyer-Whiteford points to a more recent current, referred to as autonomist Marxism, that expands the limits of traditional Marxism and works to inventively reassess the foundations of Marx's work and extend it to contemporary world. He ascribes to this line of thought authors such as Tronti and Negri. For Dyer-Whiteford, this current is marked by the very fundamental perception that capital has had a large centrality in Marxist theory, as an inexorable organizing force, while the working force had been seen as passive. Autonomists believe that this vision should be inversed: "the worker is in fact the *active* subject of production, the wellspring of the skills, innovation and collaboration on which capital depends." (p.65) Therefore the use of the term "autonomy".

They envisage three main cycles in capitalism, regarding working class struggles, that correspond to different strategies of control and to different forms of resistance: the cycle of the professional worker, of the factory worker and of the socialized worker. The first would be the qualified artisan in the workshops, the second, would be the deskilled worker in the fordist chain of production, that shattered collective organization, and the third would be the socialized worker, where every level of society becomes a part of the productive cycle and is subject to multiple control strategies. In each cycle, technology is used in different ways for organizing living labor and to seek a relative emancipation of capital. According to Whiteford, contrary to the idea that the workforce in capitalism is seen as a mass, capital draws different social groups for better social control on a global scale, but at the same time develops several strategies to break apart and deskill collectivities that resist its exploitation. So the working class is defined by its struggle and by the living labor. Contrary to the Post-Marxist critique, autonomists will insist that this definition of the working class does not have to deny the diversity that is present in this struggle, and that, on the opposite, it will mess up the categories in which capital seeks to place workers for better control.

Moreover, they understand that working class struggle is precisely what drives capitalist development: "capitalism does not unfold according to a self-contained logic, spinning new technologies and organizations out of its own body." (p. 66-67) In this sense, while capital strives to develop technology as some sort do dead capital in order to better control the living working force, or even try to emancipate itself from it, workers use their creative power, the same that capital needs for innovation, as a form of reinvention and appropriation of

technologies. For the socialized worker, it is especially the communication technologies that are reappropriated. At the same time that capital extends its control beyond the factory, and into education and private life, integrating the society as a whole as part of capitalist production and consumption cycles, different collectivities are recomposed. It is an emerging "worker's use of science" (p.71).

So in Marxist theory, especially in the recent work of autonomists and more clearly from the context of the socialized worker, we see the span of political action extending far beyond that of the State, and far beyond its organized arenas of debate. Indeed, according to Virno (2008), with digital communication and information technologies, and considering their development as part of this dynamics of emancipation and resistance, there is a change in the traditional relationship between political action and the two spheres that traced its perimeter: work and intellect. Work, for being penetrated by innovation as its *modus operandi*, loses some of its repetitive character and becomes a space for constant reorganization and negotiation, clearly displaying the multiple roles technology plays in social struggles. Intellect, on its turn, instead of something of an individual level, abstract and invisible, becomes visible in multiple communication and information technologies, and takes central stage in drawing together the collectivity. As systems of capitalist management become more pervasive, so does the political sense of every social interaction, so political action penetrates work and intellectual activity.

Related to this we have the discussion proposed by DiSalvo (2012), who brings up the idea of agonism, used to describe the passion underlying public debate and the conflictual character of collective life, and relates it to its possible embodiments in several technological devices. According to him, "democracy is not simply order and rationality displayed in voting, structured decision making, and legislating, but also and necessarily is contentious affect and expression" (p.4). Our heritage since the Enlightenment and the ideal of the modern democratic state seeks consensus through debate, but agonism brings the idea that social life is in itself a condition of disagreement and confrontation, of pluralism, of constant reorganization and controversy that cannot be reduced to democratic institutions and the staging of predefined positions.

So we should be able to differentiate politics from the political, which will help us differentiate practices from institutions. DiSalvo refers to Mouffe when he explains that politics are an "ensemble of practices, discourses and institutions which seek to establish a

certain order and organize human coexistence" (p.7). Politics is what Habermas would call the institutionalization of the public sphere. From a different perspective, we could also add that they are the structures of management of social life. Politics embody certain representations of social life and, regardless of the many stable procedures developed, are always potentially conflictual because they are affected by the political debates in general. The political, according to DiSalvo, is a condition of life. It is inherent to social interactions and can take many forms, since living together will always fundamentally entail negotiation and debate. This transfers back to the public scene and therefore to public debate their fundamental political role.

So we have this double sided problem which is the coupling of the limitations of *representative* democracy, that institutionalizes debate with the alignment of institutions and generalizing positions, with limitations in the *representation* of the actual public issues, that is far too reduced in diversity of variables and arguments, and, as Gomes puts it, stages social struggles from the outside. At the same time, we have the notion that social struggle is pervasive, and that there is a collective intellect embodied in devices whose struggles become apparent in technological development. So this concern with representation becomes deeper once we understand that the problem of representation of social issues must incorporate the conflictual and agonistic character of technological development and appropriation themselves.

3.2. MANAGING SOCIAL LIFE

Foucault describes a long set of transformations in the attitude towards sexuality from the eighteenth century on, where first there would be a contrast between the way it was experienced in the countryside and in the new disciplinary spaces that were being structured in the new urban centers, followed by an association between scientific inquiries on behavior and health, law, criminology, public administration and institutionalized education, to develop boundaries between the normal and a myriad of deviants and to promote an economy of private life and of bodily behavior. Likewise, as commercial exchanges extend over larger and larger territories, and cities start to grow, different challenges for managing all these many commercial flows and other activities, and also flows of people, of garbage and merchandise, they demand monitoring, records, values, rules, laws. Even as private property and private spaces are reinforced, there is the complementary public management of private behavior.

All those related movements have in records and statistics some fundamental tools. Tarde (1883), for example, by the end of the nineteenth century, compared the statistical curves in the quantitative graphics to the line suggested by flight of a swallow. While statistics turn fluctuations into appreciable values, the graphics return their continuity so their aggregated sense may be concatenated and narrated. The same way as the eye sees many different states of the movement of the swallow as a continuous line. He moved ahead by proposing that in the future newspapers, instead of publishing polemics in literary form for the exchange between the very well-read elite, would give statistical graphics a place of honour. Newspapers would therefore become the publishing terminals of several statistics offices, constituting the social equivalent of the organs of the senses (p.509). Even though Tarde does express some concern for the ways in which statistics are used, these ideas show a very strong disposition to consider statistics (and data in general) as the most reliable access to reality, and the graphics themselves as the devices for access.

As the twentieth century approaches, the relation between the public management of the private life of the masses and the management of public opinion of the general public only becomes narrower. Statistics are applied in various fields outside economics, and used as an instrument for establishing social groups and developing knowledge about them. So the use of data and statistics to manage public opinion only deepens in the twentieth century, be it in opinion polls, voting surveys, demographics, audience rates, consuming patterns and others. This represents the emergence of a complex set of disciplines to deal with the deprivatized publics, whose profiles and personal choices become something of great importance for institutions of private interest and for governments. These disciplines, as ways of describing collectivity, work, from end to end, by structuring political positions, market shares, demographic groups and so on. Knowledge is generated, from the beginning, though classifications that create a categorical map of society in order to organize it for different strategies.

Hacking (1995) also understands that this kind of institutionalized classification of people was possible only from the emergence of the industrial bureaucratic systems, in the first urban centers. He refers to this phenomenon under the name of *human kinds*, and seeks to identify some of the criteria for their constitution and evolution, as well as some the effects these kinds may have over the people that are classified. According to him, “human kinds are formulated in the hope of immediate or future interventions in the lives of individual human beings.” (p.351). They are instruments for state of the art research and for social intervention, for

framing and understanding phenomena and making reliable predictions. This will include of course the description of demographic characteristics, but also and specially the causal relations between certain behaviors or social conditions to measurable variables.

These human kinds are used by many institutions and specially the media, to refer to and describe social groups affected or involved in current events. They reverberate to wider publics and are incorporated in public debate. This happens whether the kinds are based on race, behavior or consumption patterns: human kinds start by one point and incorporate other dimensions as the profiling work advances. This classificatory vocabulary gets to be adopted by the same ones that are classified and has a strong adherence to the discourses about collective issues and in political action and social engagement in general. Members of a certain human kind tend to identify themselves as such, adhering with more or less restrictions to associated discourses and demands, and also producing their own. For Hacking, identifying and naming these kinds has the effect of grouping and levelling the people under them for social intervention, but at the same time opens the possibility for looping effects, that is, of mutual constitution, when, for example, the classified persons assume their kind as a discursive tool for passing demands and engaging collectivities. Therefore, human kinds, contrary to human natures, bring on constant translations and changes in the criteria of use and for behavior prediction, and their interpretation will vary in time with the developments in the state of the art of research and policy, and with the social tensions at stake, having to be constantly reevaluated. Like the stereotypes that drew the concern of Lippmann (1922), the human kinds also become landmarks defining the social places of different actors, but Hacking considers from the start the work of manufacture involved as a complex set of collective study, practice and behavior, and signals to a deeper interaction between many different actors when he points to the looping effects. Those actors can be scientists and scientific institutions, politicians and government policy makers and administrators, companies, product managers and designers, concerned publics and consumers and so on: human kinds are translated and appropriated, and do not belong to specific actors, they are in the wider sphere of the public scene.

Recently, the penetration of innovation in the work environment has to be coupled with highly skilled and productive workforce that needs science, communication and the communication of knowledge as its raw material. Dyer-Whiteford (1999, p.83-87) describes a schematic opposition between communication and information, as proposed by Negri: information would be similar to dead labor, part of inert representations of reality, once communication

has been expropriated from its actors. It is hierarchic and centralized in its structuring. On the other hand, communication, like living labor, is inventive, transversal exchange. Once more, emphasis is given to living labor as the autonomous force at the origin of information. The repurposing of information technologies reverses "the cycle of information into a collective organisation of knowledge and language." (Dyer-Whiteford, 1999: p.71).

We could add that, inside as well as outside of the work environment, the management of information flows becomes critical for work, political action and collectivity. It is through communication that information is not only produced, but also reassimilated into knowledge. It is through communication (or exchange, public debate) that information may be reconducted as part of political action. Nevertheless, the effects of the accumulation of these flows and of the creation of different interfaces (as contexts for displaying information that are sometimes incompatible between each other), create a new level of challenge for broad communication and the distributed production of knowledge, because they stretch human cognitive availability to a limit. For Boullier (2012), we should keep in mind that this availability is not just in terms of time (already scarce), but also in terms of the spirit, of the attention that is requested by informations and attractions of all kinds. To put it briefly, he understands this to be a cognitive issue.

3.3. ATTENTION ECONOMY, ONLINE MEDIA AND THE PUBLICIZATION OF THE INTELLECT

As social interactions, in and outside the state and institutions, are very distributed, it is hard to keep track, fully debate and develop a cohesive understanding of social reality and its events, so people rely on specialized translators like journalists and delegate wider debates and decisions to politicians. But this, on the other hand, bounces back as more competition for attention, because politicians will seek popular support for their positions and proposals (especially but not exclusively in electoral periods), and media outlets will compete for audience. What Boullier highlights is not exactly the issue of the excess of information produced nowadays, that is so dear to many discussions about visualization, but the issue of the excess of requests that fractures the continuity of our processes of attention, affects our concentration, and hinders our possibilities for developing our own cognitive syntheses. This proposes a fundamental shift in our discussion, because it brings up the concern for the effects of an *economy of attention*. Citton (2013) understands that this economy of attention is not

new and that, to an extreme, there are problems connected to it in all communicational processes: there are always discursive strategies being developed so that a certain message gets the due attention, even in informal conversations. For the author, what is important is understanding that, in the information age, the main rarity that will be scarce and has to be carefully managed and sought after will be the human attention, the cognitive availability that is necessary to make sense of the abundant information available. So many practices for the management of attention come into play, as the ones from media outlets, for example, also from marketing and advertising, and a host of other strategies to which we are exposed as we go on in our everyday lives.

Boullier (2008) ponders that composing a common world becomes, therefore, a much uncertain task, while everyone goes on with their occupations and preoccupations, skipping from piece to piece in a fractured landscape. For him, we would do good to rediscover the dynamics of attention of traditional societies, which kept a sense of care and continuous dedication that would allow us to recapture a relation to the cosmos, or at least to reconnect and deepen our experiences. In opposition to the current attention regimes, this regime of lasting attention that the author associates with traditional societies would be formed by a capacity for taking the time to pay attention, for a long duration, to all the surroundings, that constitutes a form of caring that is delivered to all the beings in the cosmos. Nowadays, we would have fractured and superficial attention at work in a context of struggle with the many demands for the affiliation of the general public. Curiously, we would tend to relate the concentration of lasting attention to crafts that demand individual focus, like writing or manual work, but for the author it shows up as having the force to reconnect social relations. This can be related to the cycles of attentive recognition as described by Deleuze (1990), that we explored in the previous chapter: deepening the understanding about something by reassessing it many times, while at each cycle reconnecting it to a wider cognitive reserve, broadening and integrating the landscape.

So in fact what we call *cognitive overflow* (KIRSH, 2000) would not be caused exactly by an excess of information, but by a competition between attractors that does not allow this dedicated time for caring for and deepening our relations with the world, which would at last aggregate a cosmos and an integrated society. Regardless of the amount of records produced, essentially the amount of information that is assimilated always depends on processes of attention. Maybe the difference is that, in the traditional societies described by Boullier (2012), in order to have more information, it was necessary to observe more and for longer

periods of time. Meanwhile, nowadays, it is with the superficial and fractured attention, that does not dwell deeper, that one has contact with more records (taken as equivalents to more information), along which he/she navigates generically. Even though the amount of records is an issue, we recognize once more that the fragmentary character of this experience is the key.

So the data and information deluge does not worry only because of the excess, but because being excessive means it becomes harder to create contexts for cognitive syntheses and for exchange, the roadmaps. On the other hand, social phenomena are not always self-evident, because they are composed of many distributed interactions, so even experiences with social events that have a direct physical existence might need some effort in order to be put into context and being suitably interpreted. So, to some extent, associating things to visible records becomes a condition to their social existence. Nevertheless, once everything is potentially recordable, visibility can turn into its opposite, it can obscure understanding. So we would like to advance that there is a change in this demand for visibility, especially from the second half of the twentieth century on: more than uncovering and storing things, it becomes necessary to build devices for relating and comparing records, to contextualize the access to them. So visualization comes to sort out a problem of visibility but also of mediation.

Digital networked media does not exactly the create a space that did not previously exist, but equips public debate with new tools. From our point of view, this equipping goes in two senses: first, it displays more clearly – or we could say publicizes – varied points of view and processes of identification between people, institutions, products, ideas and other people, in such a way that keeps and amplifies most of their fluctuations and plurality. The attributes of these processes do not emerge with online communications, since public opinion has always been the fruit of much noise and plurality, despite all the efforts for its management. Nevertheless, they become more evident thereafter, and can be tracked in their processes and not only as stabilized relations, when interactions are recorded and processed infinitely. What happens is that an enormous field of commentary and exchange becomes more visible and measurable, even through and inside the strategies of institutions of private interest and the state.

We borrow from Berger & Luckman (2005) the idea that collective debate, information exchange and, at an extreme, social acceptance, produce objective knowledge, and not the other way around. With that we do not mean to argue that every knowledge is a fabrication or a staging for other interests, but that they only become objective after much work and

discussion has been invested in making them so, after it has been connected and asserted by the performances defined in institutions of disciplinary knowledge and social authority. Objective knowledge demands methods, verifications, institutions and so on, but it also demands public concern and engagement, so that it can be more shared and accepted. So, fundamentally, more than access to information, debate itself is at stake for contemporary democracies, for turning information into shared knowledge. We propose that visualization can be a useful tool to that end.

To better develop our conception of visualization as a device for collective and social performance, we bring up the relation between work and intellect, as traced by Paolo Virno (2008), a philosopher of marxist influence, that Whiteford includes in the group of the autonomists. While seeking to address the problem of political action in contemporary societies, he advances that, according to traditional political theory, the idea of action could be outlined by two frontiers: at the one side, there would be work, and all its morose nature, and the automatism that makes it a repetitive and predictable process. At the other side, there would be intellect, or pure reason, and its solitary and unseen nature. Political action, therefore, would be different from work, because it would intervene on social relations rather than on raw materials, and would be related with new and unforeseen possibilities. Also, contrary to intellectual activity, political action is public, contingent. Virno posits that we can no longer rely on these differentiations between action, work and intellect: first, because work has absorbed traits of political action, and is carried out through constant variation, adaptation, negotiation. Second, that the intellect has become public, has broken into the world of appearances by a coupling between private reasoning and information tools, becoming also a key for collective life in an update of the notion of *general intellect*, as proposed by Marx. On both sides, we see the problematization of the embodiment of discursive performances. We believe that it is in this sense that Disalvo (2013) proposes the design of contestational objects, as a performative embodiment of political action.

According to Virno, Marx considers the general intellect as the objectification of the natural sciences in productive machines, as a form of fixed capital. Virno, on the other hand, proposes that we give maximum prominence to the general intellect as a direct attribute of work, a repertoire of a diffused and collective intelligence that gathers a crowd. What interests us most in this discussion is the idea that with the publicization of the intellect, the human faculty of reasoning for reason itself (KANT, 2005) gets translated into the most basic cognitive capabilities of the human being, that are the condition of every composition and of

every experience, and are now coupled with information and communication devices. So the most fundamental and generic features of the mind, like the faculty of language, the disposition for learning, the capacity of abstraction and correlation and the access to self-reflexion, become public resources that gather collectivities, we should add, by information exchange and debate. In this sense, we should ask how the basic cognitive capability of pattern recognition, as promoted by visualization, could be part of this shared repertoire for composing spaces for collective action.

On the other hand, this signals to the concern for the structures that are going to organize the use of these basic repertoires, and how they may channel and set the stages for what can be said and how should it be said, creating frameworks for the general intellect. Manovich (2012) argues that there is an infinity of different architectures and devices that specify "how information is presented visually and/or spatially, how it is updated over time, and how users can interact with it". He calls them *design patterns*: these are deeply influential to how we experience information and, in a heavily mediated landscape, with how the publicization of the intellect will turn into social practices.

3.4. DEBATE, MEDIATION AND ACCUMULATION

According to the british empiricist Lippmann (1922), social reality is out of sight: it is too big, varied, complex and fleeting to be fully understood, so most of the time it should be imagined. And this, according to the author, leads to a key problem for modern societies, that is the fact that people will often build distorted versions of reality, not only caused by emotional issues, but because of the stereotypes that are built as interpretive tools that – like the mental schemas we have been discussing in the previous chapter – carry previous interpretations. They tend to build a *pseudo environment* in which people act and that would determine a great deal of political behavior. These stereotypes start from personal experience, but are really reinforced and structured through interactions with others. Groups of people that share similar references use these shared projections to position themselves in the group and draw their appreciation of the positions of the others and of outsiders, and their value. We can say that, within these groups, we have the creation of systems of stereotypes that are developed into frames of reference, that organize what should be discussed and how. Stereotypes are stabilized by practices, and will also need some suspension of critical thought. A related and perhaps even

more serious concern is the fact that these stereotypes, that are in the core of the building of public opinion, can be manufactured by different competing forces.

To address both issues, Lippmann proposes that journalists should act as specialized translators of social reality, so that citizens could have access to reliable information, in order to develop their opinions and transcend their stereotypes. One problem with this solution would be that the specialists themselves, despite devoting their lives to suitably reporting on social reality, also have stereotypes in their minds, affecting their judgment of the facts they are collecting, and can also be influenced by many interests. Nevertheless, we believe Lippmann's proposition, after all, has roots in the more fundamental idea that the circulation of information and the broad public debate are the most efficient tools for counteracting the disadvantages of the pictures in our heads, the stereotypes, and to dismantle much of the strategies for their manufacture. Moreover, we also propose that broader exchange of information develops the conditions for a shared world, if it manages to cross barriers of those systems of stereotypes and develop a wider and shared plane of reference for a more inclusive debate. In this sense, and also considering the important and even structural role that data has taken up in heavily mediated societies, we believe visualization could collaborate in building these common spaces of debate towards wider and more inclusive representations of social life.

Nevertheless, while understanding the important role that media (especially journalism) has for the broader circulation of information in modern western societies, we must carefully consider the specific challenges their content strategies and methods pose to allowing richer and more inclusive debate. According to Wolf (1987) in his work on communication theory and mass media, with the progressive mediatization of social relations by technical devices from the first half of the twentieth century on, there is a tendency of stressing the efficiency and clarity of communicational channels, while communication itself is approached more as a process of transmission than of sharing information. He indicates a communicational model derived from information theory and from the fundamental works of Nyquist, Hartley and Shannon, that form mostly a theory for informational efficiency. As we are going to discuss in the third chapter, this theory sets aside issues related to the meaning of messages and puts processes of transmission of information between people, between people and machines and even only between machines at the same operational level. The central issues gather around channel efficiency and appropriate encoding at one end, transmission with minimum loss and decoding at the other end, through a uniform code.

Wolf also points out the importance of the work of Jakobson in the integration of linguistics with the mathematical theory of communication, leading to the broadening of the idea of information, and also to its abstraction. As we previously discussed, information becomes that which remains constant across all the reversible operations of translation and codification. Therefore you have, at the same time, a wider and less specific scope for information theory, and a communicational model that is focused on a functional relationship between sender and receiver, where the code is to be considered uniform and the reception is restricted to matching the literal sense of the message on an individual level.

Nevertheless, with the development of semiotic studies and sociology applied to communicational problems, there is a growing awareness of, on the one hand, the meaning of messages as derived from the communication processes and, on the other hand, of the collective and cumulative effects of information flows. Wolf names this, *apud* Eco and Fabbri, the semiotic-informational communicational model: communication is seen as the transformation of one system by another, due to semantic exchange introduced by codifications. This model will assess communicational processes by their dynamics rather than by their efficiency, and will highlight the processes by which the public builds up meaning from the messages received.

As this model is mostly confined to analysing messages, their codes and communicational structure, Wolf advances that it is not fully equipped to deal with larger communicational phenomena, such as mass media. He describes then a third communicational model, the semiotic-textual: the general attention for individualized messages is replaced by a concern for sets of textual practices. According to him, with mass media, people do not only receive messages that are recognizable according to known codes, but sets of textual practices, that will compose the basis for the recognition of grammatical systems for many messages.

In that sense, this new perspective, associated to the sociology of knowledge, will progressively build upon the understanding that communications do not influence directly the explicit behavior, but are the means by which individuals will organize their image or perception of the environment. So there are some major changes in the methods and procedures of communications studies: instead of focusing on specific cases, like campaigns, news series or even messages as individualized objects, there is the tendency to focus on specific themes and trying to understand their representation across several medias; less use of

survey data and the development of complex methodologies based on other kinds of data and traces; less interpretation on the changes of opinions and attitudes and more effort on reconstructing the processes by which the individual modifies his/her representation of social reality (NEUMANN:1983, *apud* WOLF, 1987).

Wolf draws attention to the fact that this general concern for the effects of communication and mediation processes in the social construction of reality had already been evident before recent years, although in a more fragmented way. Walter Lippmann (1922), for example, shares this concern for the effects of communications, and describes the pseudo-environment that everyone builds from the information they receive, the stereotypes, or *pictures in our heads*, and the difficulties of representing social reality. We should note that this new paradigm brings the research perspective of communication studies quite close to the general perspective of controversy mapping, because of this passage from the analysis of the message as an object to the analysis of meaningful exchanges as part of the formation of a social context.

Hall (1993), while working on the more specific subject of the television medium, discusses yet other limitations of the traditional communicational model of the sender/message/receiver, and proposes other forms of describing and approaching the communicational process, that may reconcile the message and its form with the collective aspect of its circulation, and the many translations involved. First, he points out to the fact that the traditional model, understood as a loop, is too linear and restricted to the transfer of the message, without a solid conceptualization that would account for the complex set of relations that is present at each moment, at each translation. He proposes the articulation of four distinct and related moments: production, circulation, distribution/consumption and reproduction. Each of these would represent a set of interconnected practices that influence the other moments without fully defining them. With this, at the same time he addresses and associates, on the one hand, the productive and material means of communication, and, on the other hand, the discursive form, where the circulation and the distribution take place. He argues that the operations of meaning are the fundamental moving force: “Once accomplished, the discourse must then be translated – transformed again – into social practices if the circuit is to be both completed and effective. If no ‘meaning’ is taken, there can be no ‘consumption’. If the meaning is not articulated in practice, it has no effect.” (p.91).

So the moments of encoding and decoding become determinant not only for the continuation of transmission from one end to the other, but in the attribution of meaning and, therefore, to enable the complex set of relations in each moment and in the passage between them. A social event, when reported, must pass under the sign of discourse, it "must become a story before it can become a *communicative event*. (...) The 'message form' is the necessary 'form of appearance' of the event in its passage from source to receiver." (p.92). This moment when the event is set in a discursive form involves the dominance of the rules of discourse, set themselves in an organization of social relations. According to Hall, this leads us to the realization that:

> "circulation and reception are, indeed, 'moments' of the production process in television and are reincorporated, via a number of skewed and structured 'feedbacks', into the production process itself. The consumption or reception of the television message is thus also itself a 'moment' of the production process in its larger sense, though the latter is 'predominant' because it is the 'point of departure for the 'realization' of the message." (p.92-93)

So reception incorporates the message into society and integrates its perceived meanings in world views and behaviors, that will also be reinterpreted as social events by the media. We should draw attention to the fact that these feedbacks are *skewed*, because, for Hall, the determinant moments of encoding and decoding may not use perfectly symmetrical codes. In one moment, the media infrastructure applies a code and forms a message, and in another moment the message is decoded and reenters social reality in the form of social practices, to be reinterpreted by new mediatic messages. Between both moments there are many differences of position, once the relation between the public and the medias is not symmetrical and the situation of those moments differs greatly.

So we are back to the issue that brought much concern to authors like Lippmann: according to Hall, the knowledge that derives from these many discourses will not be a transparent representation of the real in the form of language, but "the articulation of language in real relations and conditions." (p.95) Hall also posits that there is no transparent code: the effect of realism or fidelity between the representation and the thing to which it refers is the result of a discursive practice, of a "certain articulation of the language with the 'real'" (p.95). Therefore, the possibility of a closer alignment between the codes used in encoding and decoding moments, in situations where a code is widely and much uniformly distributed in a society, does not point to a more natural, transparent or realistic representation, but to the depth, the deeply rooted habits in which this dominant code is set. Stabilized codes like this

carry the dangerous effect of concealing ideological practices of coding that are always present. An equivalence of codes that is produced by habit may, in a broader communicational level, impoverish the exchanges of meaning.

One last important point that is made by Hall refers to the difference between denotation and connotation in communicative processes. He draws our attention to the fact that "denotation" is often considered to be a stricter or literal translation of the reality to which the term may refer to, but is, in fact, a more strict and stable connection to a specific *meaning*. "Connotation", on the other hand, is associated to a less rigid and more conversational use of words, through associative meanings, and may vary according to the situations of use. While the connotation of words displays more clearly the interference of ideologies or interests, the naturalization of denotation may hide these same ideological meanings, that are equally present. The main difference is that in denotation, these ideological contents are more fixed, stabilized. So both categories, connotation and denotation, should be used more in analytical terms, as tools for distinguishing different levels in which ideologies and discourses intersect.

What is most interesting to us is that, from the appreciation of these two categories, Hall proposes a very powerful interpretation for the role of codes to developing certain relations between signs and the wider social and ideological universe. By connotation, "the environmental world invades the linguistic and semantic system" (BARTHES *apud* HALL, 1993). Hall advances that codes are the means by which power and ideology are continued, reinforced or challenged, and are led into social practices by driving meaning to particular discourses. Socially shared codes create references between signs and *maps of meaning* into which culture is classified. Hall actually describes codes as *maps of social reality* that "have the whole range of social meanings, practices, and usages, power and interest 'written into' them" (p.98).

These maps are always the object of much dispute. However changing and polysemic the many socially shared codes may be, any culture tends to impose its classifications of the social, cultural and political world in a *dominant cultural order* that organizes different discursive spaces of dominant meanings. Every new event, however disruptive of this order it may be, will have to be assigned to a certain discursive domain in order to make sense in general terms, which paradoxically may drain much of its significance. So, according to Hall, this assignment mobilizes a set of *performative rules*, that develop competences and uses, in

order to reinforce one semantic domain over the other at each situation, and move items from one setting to another, transforming their affiliation and meaning. So the reception of communicational messages and the integration of them into actions has this deep and fundamental performative aspect that will, of course, be reintegrated in the skewed feedbacks of reproduction.

We believe that the language and the boundaries of what can be said and how it will be said, this discursive practice, are the main site of dispute in communicational processes. We should point out that the skewed feedbacks described by Hall are equally important outside the dynamics related to the television medium and even to mass media. Online media is a very fertile field for these disputes in encoding and decoding, where, by a new distribution of publishing devices or communication infrastructures, the polysemy of codes and translations is publicized and therefore becomes more visible. Of course, the dispute between different codes and different meaning attributions is still present, but accumulated with other strategies. As an example, we have the work of professional media outlets towards increasing the penetration and visibility of their content in social media. As we have discussed elsewhere (CARVALHO & BOECHAT, 2015), cartographies of social media activity may display quite clearly these disputes of credibility and relevance and the concentration of flows around these major social actors. To a certain extent, this configuration reproduces the dynamics of mass media, as described by Hall, but, from another point of view, this new environment keeps records and amplifies the possibilities of other minor flows, that may display the effects of the appropriation of these dominant codes to the social organization of meaning.

Of course, in a media environment dominated by mass media, the major flows are more visible, but, at the reception and consumption moments, with the processes of decoding, there has always been much circulation and skewed reproduction of messages. The problem was that it was not visible on a collective scale, it was hidden under the major flows, so dominant codes were more easily regarded as transparent, their strategies hidden in plain sight. As Wolf (1987) explains, contemporary strands of communication theory will seek in semiotics and in the sociology of knowledge the tools to overcome an approach that focuses on understanding the possibilities for matching and uniform interpretation of dominant codes and will dive into those disputes along encoding and decoding moments.

In journalism theory more specifically, some of the main forces at play in these disputes are discussed in the theory of agenda setting. According to it, news outlets influence

the formation of a public agenda, which is nothing more, nothing less, than a map of issues and the corresponding code in which enunciation and reception might take place.

According to portuguese journalist and researcher Traquina (2002), agendas are the set of current affairs being followed in different spheres of society. It can be a mediatic agenda, referring to the discussions presented in the media outlets, an agenda of governmental policy, meaning how the state translates different agendas into policy, and a public agenda, meaning the issues emerging from the direct interactions and opinions of common citizens. They can be very different, both in terms of priorities and in terms of how the issues are framed, and they also exert pressure to affect one-another. Journalists, for example, may frame events in such a way that has little to do with the issues of interest of their public, it would be reasonable to imagine that media outlets may define the public agenda. Nevertheless, for Traquina, it is not so straightforward: even though the influence is clear, if a certain mediatic agenda has no resonance with the concerns or references of the public, it will not be adopted. On the other hand, newsrooms are also like boxes of resonance of the many pressures issued from state and private institutions, that will willingly provide overworked journalists with information and approaches in accordance with their interests, with the intention of influencing public agenda through mediatic agenda. So journalistic practice itself is affected by strategies of agenda setting, if not only from representing and reencoding the effects of journalistic agenda setting on public debate, also by being the target of the strategies of other institutions and their respective agendas.

Traquina also refers to Molotch & Lester, who defined three kinds of actors that take part in these dynamics, from the point of view of journalistic practice: the news promoters, people who identify an occurrence as relevant (therefore making it observable) and exert pressure and feed reporting; the news assemblers, professionals who transform an observable and finite set of occurrences into public events through publication therefore defining mediatic agenda; and the news consumers, who access certain occurrences that are made available by the media outlets. The news promoters, in this classification, are specialists, people who propose the governmental agenda, people from politics. It is quite clear that with digital networked media the news promoters also publish content that participate in the public agenda, and more people from the general public are included in this group. And, of course, news assemblers are also consumers.

Before digital networks, journalistic agenda setting seemed to place itself outside of public debate: for Lippmann (1922), journalists should be the experts that would be able to identify the issues in the public agenda from the outside, and search for comprehensible and true information to feed back to the general public. As public agenda becomes more visible, like in the social medias, for example, journalistic practice is more deeply implicated with these translations between different actors. These exchanges point to a process of construction of public issues through journalism: while an occurrence is just something that happens in time and can be visible, the constitution of a journalistic event means giving it public existence and constituting it as a resource for debate. Then, the biggest concern is that journalists get isolated from public debate and too close to governmental or institutional agendas, and may be framing events in a way that hides relevant issues. In order to avoid this problem, the solution in journalistic practice in general would be to search for as many voices as possible (DURHAM, 1998), so the issue bears as much grounding as possible to actual controversies.

At the same time, Wolf (1987) draws attention to this cumulative effect of mass media: the capacity of mass media to create and maintain the relevance of a certain theme is the accumulated result of the workings of media coverage, of repeated exposition, and not isolated effects. And this has continued effects, because even if different informations are given on the fringes of the main accumulation, the common features and similarities tend to be taken as more significative than the different approaches.

Hall, on the other hand, refers to the idea of *selective perception*, that is used by researchers interested in understanding the deviant interpretations of media messages on the part of the public, and how to reduce them. According to Wolf, selective perception is understood as a very specific way in which individuals decode messages: according to how attached they are to their values or previous assumptions, they may chose to highlight and de-contextualize certain aspects of a message while toning down others, or the whole significance of the message. Nevertheless, Hall relates the use of this notion mainly to the concern that the audience might fail to grasp the meaning of the message as intended by broadcasters. He sees this notion from a more productive point of view: “This [selective perception] is the door via which a residual pluralism evades the compulsions of a highly structured, asymmetrical and non-equivalent process.” (p.100)

Considering the comulative effect of media messages indicated by Wolf and adding to it the information deluge we face in contemporary networked societies, we can see that this selective perception may be a *defensive* perception that not only defends a person's values, but forces many messages into simplified models, for the sake of maximizing cognitive capacity quantitatively. We relate this tendency to the cognitive efforts of, when overexposed to information, trying to recognize patterns and to develop standard schemas to connect with, directly in action-reaction loops, as we described on the first chapter.

We should also note that Wolf talks about a process of agenda setting by accumulation, of feeding large amounts of information according to the themes and approaches present in the desired agenda while leaving other themes, considered to be less relevant, to be poorly represented in terms of occurrence and detail. These last, despite present in general news coverage, get, as time passes, cumulatively buried by the rest, so we might also talk about agenda setting by superposition, or even occlusion. This is also a fundamental feature of the specific sort of economy of attention we live inside a context of information and cognitive overload: more than the traditional competition of several attractors for their messages, there is an accumulation organized by dominant codes that gives form to excluding communication landscapes.

In this sense there is a capacity of pattern recognition being developed, but it is a defensive one, much different from the pattern recognition that is associated with the exploration and interpretation of visualizations, because it works by stabilizing codes and mappings. As we described in the first chapter, it is the continuation of the image into utilitarian action by the sensorimotor system, in direct translations of automatic cognitive reproduction. This does not mean that the codes in the encoding and decoding moments will necessarily match, but mostly that they will configure systems that do not communicate in a deeper and transformative sense. Selectivity and occluding accumulation will empty most of the informative quality of messages, in the sense that they will no longer be capable of informing or being actually translated into other systems of attribution or new practices. The production of forms that is connected to the circulation of messages and information might be varied, but will also be very superficial. So there is, in communicational processes and discursive practices, effects that have been described at ANT as blackboxing: gathering complexity in simplified representations that fade into the background of tacit assumptions (LATOUR, 2009). In the case at hand, this creates the impression of transparent

representation, and also reproduces and reinforces the separations of discursive spaces that are present in every society.

We believe that visualization can help to counter that tendency, not precisely for dealing with large amounts of information, but for three main reasons, that we intend to explore in the rest of this thesis: first, it may compose shareable planes of reference, where categories are found, can be built from the bottom up, and not necessarily imposed by previous schemas. Second, as we signaled in the previous chapter, it can build different forms of access to information, integrating the processes of display and exploration with the attentive recognition, integrating scattered records into continuous landscapes to counteract some of the effects of the contemporary economy of attention. Third, for transcending individual cognitive aspects and becoming tools for collective work, mediating devices around which people can gather.

4. THIRD CHAPTER: DISCRETIZATION AND VISUALIZATION

Over the previous pages we have been dodging a central issue to this thesis and it is time we deal with it head first: what is information, what is data and how to pull them apart, if ever? The task can become surprisingly challenging, considering the intricacies of technical mediation involved in the development of both categories. Information designer Zer-Aviv (2014) offers us a very strong and streamlined definition of information and data, summarized in the figure 3.1.: the DIKW pyramid displays the progressive relation of data with information, towards knowledge and finally wisdom. He explains it is

> "a framework for thinking about how data gains context and meaning and becomes information. This information needs to be consumed and understood to become knowledge. And finally when knowledge influences our insights and our decision making about the future it becomes wisdom. Data visualization is one of the ways to push data up the pyramid towards wisdom in order to affect our actions and decisions."

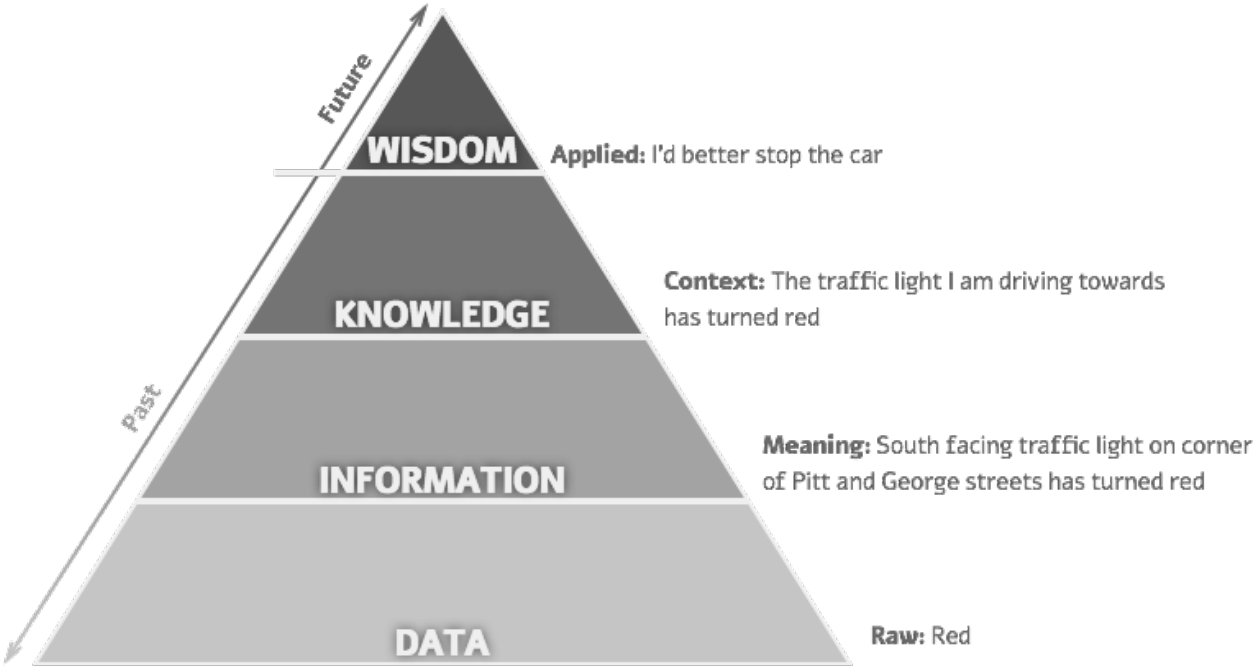


Figure 4.1. DIKW pyramid, by Zer-Aviv. Source: https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis.

Despite drawing attention to several issues of our interest, like the progressive chain of translations from data to wisdom and the relation between context and knowledge, his formulation is closely connected to decision making and a linear temporality, associated to the continuation of perception into useful action, like the sensorimotor schemas of action and

reaction described by Deleuze (1990). It does not account for the collective and social aspects of either knowledge or wisdom, that are central to our discussion, neither for the coupling between information, technical devices and visual display that provides the base for these movements. We will explore this second theme in this chapter, along the thread of a progressive discretization of languages, supports and technologies of data and information, always accompanied with complementary movements of reassemblage.

Gleick (2012), while building a history of information, advances that every new technology reorganizes the world around it and forces people to deal with new facts while adapting old terms to describe them, which leads, at first, to much inconsistent meanings (p.161). We have witnessed this phenomenon in the last few decades with the many visual and verbal metaphors that rapidly populated the emerging digital communication technologies, like the term "e-mail" or the desktop in operational systems, that refer back to previous devices. At times, in software design these inheritances can be used on purpose, as a strategy to ease the learning and adaptation of users to new environments. This is related to Gibson's theory of affordances (GIBSON, 2014): the idea that in every environment people look for clues for possible actions or points for interaction and memorize them. So, in software design, inserting elements that carry some reference to previous experiences of users will ease learning and interaction. In this case it is a design strategy. But, most of the time, these adaptations are improvised, generating much confusion, until a community of specialists develops a more or less stable understanding of the new vocabulary, usually changing the meaning of traditional terms.

Gleick (2012), while discussing this issue, describes the many changes that took place with the establishment of the first telegraph networks in the 1840's. Among other curious stories of that time, he tells the tale of a man who took a written message to the telegraph office and could not believe it had been sent because the piece of paper he had brought was still in the operator's desk (p.161). So the word "message", that referred to an object that encompassed the text and the physical material in which it was written – the written piece of paper *was* the message – started to be seen as something closer to the idea we have of information nowadays, some content that could be translated across different devices. From the piece of paper to the telegraph wires, the message was transmitted, it became something that flowed across a network of objects, being present at different places at the same time.

So the term “information” has gone through the same kind of transformation: it was distilled from the many variations in its sense, until it was a perfect neutral intersection between technological devices and practices of communication, exchange and transmission. “When it was made simple, distilled, counted in bits, information was found to be everywhere.” (p.11). So purified it became, that it is at once in the flow of every exchange and a fundamental approach for describing everything: the body is an information processor through the senses, the genes are embodied information to be transmitted and replicated, the signals transmitted by technical devices become encoded information.

Gleick develops a historical approach for discussing the role of information, pointing out relevant aspects of the interaction between mathematics, language and logic for the development of information technologies. Capurro (2014), on the other hand, privileges a philosophical and etymological point of view. He points out, first of all, that the word information has greek roots, in the term *informatio*, and that in antiquity it had two meanings: on the one hand, it was related to giving shape do some material, as in defining a form; on the other hand, it meant to transfer knowledge to someone, carrying a pedagogical and political sense. At a first glance, these two interpretations of the word seem quite different, but we understand that they are connected and even interchangeable to a certain extent: giving shape to something ascribes meaning to it, inserting it into a network of transfers that also includes people who are receiving and transferring information. Nevertheless, they point to two aspects of information: one focuses on objects or, to put it in contemporary terms, instruments and devices of information; the other focuses on information as base for knowledge, as the content of communication, available for subjective interpretation.

Capurro demonstrates that these two aspects, that he qualifies as being respectively objective and subjective, will be constantly at play in the many variations in the interpretations of the term information across the centuries. So in ancient Greece information had both objective and subjective roles, and, according to him, this continued into medieval traditions, with scholastic philosophy. In the objective sense, Thomas Aquinas (*apud* CAPURRO, 2014, p.7) draws a distinction between the physical and biological processes of reproduction of forms and the divine creation, as a transcendent cause that creates *ex-nihilo*. In the subjective or epistemological sense, he underlines the active role of the individual agent in the processes of recognition of forms that are abstracted from phenomena (*informatio sensus* and *informatio intellectus*) and also uses the term *informatio* in pedagogical and ethical contexts, that would involve wider social issues.

According to Capurro, towards Modernity the objective sense of information progressively disappears from ordinary language, reinforcing its subjective sense, of communicating something to someone. He relates this to the progressive transformation of the *substantial* man of the Middle Ages to the *communicational* man of Modernity. It also matches the emerging empiricist traditions where information ceases to be something that is underlying cosmical natural processes and becomes specifically linked to human knowledge, like an impression in the mind and the processes of memory and reasoning, and also to the moral perfection of the individual or of society as a whole. In every way, the main *praxis* of information will be the act of communicating something to someone.

From the twentieth century on, the association with the emerging information technologies, like the telegraph, the radio, and a mathematical approach towards information bring back the objective aspect of information, but with a different twist. Capurro brings up the example of Shannon's article A Mathematical Theory of Communication (SHANNON, 1948), that set the path to modern computational systems. In this article, the Bell Labs researcher detaches the idea of information from subjective meaning and from its human epistemological context, aiming at the development of discrete ways of codifying and measuring it for improving quality and reliability of transmission. For Shannon, efficient transmission of information was in focus:

> "The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design." (p. 379).

Information is converted into an engineering problem and, moreover, a matter of probability in which statistics will come to play an important role. In fact, for Shannon, information is associated with uncertainty, it implies *surprise*. If, in a transmission, only one message is possible, there is no uncertainty, no possible surprises, therefore no information. Likewise, if, after a specific letter, the probability that another specific letter follows is very high, then not so much information is conveyed. He begins to think in terms of language redundancy, of information as surprise and of messages as being sequences of discrete elements to which statistics would apply.

Shannon classifies the technologies for the transmission of information in three kinds: the discrete, like the telegraph, where the message is formed by a set of discrete symbols; the continuous, like traditional radio or television transmissions, where the message it transmitted by modulation; and the mixed ones, where there is a digital presentation of an analogical signal. Shannon's measurements and statistical approach in many ways make these different kinds compatible, something that collaborates for today's convergence in digitally based communication technologies.

Moreover, there is also a coupling with formal logic, in order to develop systems that could receive and reproduce instructions. Like Turing has done in the abstract machines he imagined as part of a few thought experiments, Shannon developed methods for encoding instructions as numbers, that were to be mapped into series of electric switches and relays, that would describe series of only two values: yes or no, closed or open circuits, electric current flowing or not, zeroes and ones. They were pairing logical operators with electric circuits; algebraic functions and machine instructions (GLEICK, 2012, p.222). This use of the term information will bring back a material and objective aspect to it, guided towards the formation and transformation of messages into networks or series of devices.

Returning to Gleick's example of the telegraph (GLEICK, 2012, p.161), we can say that information as a practice of producing forms and communicating does acquire a certain autonomy from particular objects, since the message floats between many of them, along many translations. This could lead to an emphasis in its subjective aspects. Nevertheless, at the same time, there is a complementary movement of re-objectification of information, that is reinforced by seeing it as a formal, discrete resource to be managed and, well, formatted for transmission independently of its meaning and subjective interpretations. Capurro (2014) adds that contemporary thought tends to recover the objective aspect of information because of the natural sciences, that after all must derive information from specimens, and because of the presence of information in informal language, an effect of the popularization of information technologies.

We would like to point out that one of the effects of this new objectification of information is modifying the relation between the two fundamental aspects pointed out by Capurro: the penetration of the management of information as a formal resource in wider social contexts – the political and pedagogical role of information – accelerates the translations between objective and subjective levels, inside our practices. The notion of information has always, in

a way or another, been connected to materials, but nowadays materiality is destabilized by many translations and conversions, so information is once again everywhere in the world.

We understand that there is a similar process in the uses of the term data. Spence & Wainer (in PLAYFAIR, 2005) draw attention to the work of William Petty and John Graunt, who, still in the seventeenth century, worked with many different records to manage fiscal contributions, national income and public health concerns, developing the first efforts towards statistics. According to Spence & Warner, since that time there were already large collections of data available, and analyzing them became progressively important to modern economies. Nevertheless, the existence of recorded data itself can be considered to be prior to statistics and to the maritime empires: Gleick (2012) and Bottéro (1995) point to the Babylonian clay tablets, dating approximately from 3000 BCE to 2000 BCE, that were found by the end of the nineteenth century (see figure 3.2). In cuneiform writing, these tablets displayed lists and tables of stating commercial exchanges, conversions and many calculations, and are considered to be the most ancient evidences of written language and of mathematical operations.

According to Roque (2012), in that period there was a considerable population growth in the region, which led to the development of cities and to the perfecting of techniques for the administration of common life. Conversely, the first written records are about quantities, storage of goods and commercial exchanges (p.35-36).

Bottéro explains that their writing system was not yet fully phonetic, but was already not pictorial: that was, in fact, an example where phonetic writing came about progressively, inside one same society, probably because of the intense commerce developed in the region at the time, and for the needs of communication between many cultures. He highlights the importance of these tablets as an emergent form of science, a science of the lists, that would bring to surface forms of categorization and would develop a notation removed from particular and direct experience, leading towards abstract knowledge. Gleick, on the other hand, highlights the connection between writing, mathematics and economy:

> "When scholars did learn to read the Uruk tablets, they found them to be, in their way, humdrum: civic memoranda, contracts and laws, and receipts and bills for barley, livestock, oil, reed mats, and pottery. Nothing like poetry or literature appeared in cuneiform for hundreds of years to come. The tables were the quotidiana of nascent commerce and the bureaucracy. The tablets not only recorded the commerce and the bureaucracy but, in the first place, made them possible." (GLEICK, 2102, p.45)

Roque (2012, p.39) also highlights a decisive turning point brought on by these records: the development of the concept of the numbers, despite being pushed forward by concrete and practical needs, involves some abstraction. Whereas counting things is a concrete procedure, using a single number to refer to equal quantities of distinct things is an abstract procedure.

According to Gleick (2012), it took many decades after the first excavations for mathematicians to examine the tablets and find many evidences of complex mathematical operations, comparable to geometric progressions, linear equations, quadratic equations and others. While Gleick and other authors see in these tablets evidences of these many operations and especially an early approach to algorithms, Roque (2012) will set aside this idea, demonstrating that Babylonians, even if they performed the important movement of abstraction towards numbers as we described, did not develop mathematical methods as we understand them nowadays, as linked to algebra in abstract problem solving. In their methods they were focused in calculating things that could be measured, that is, their own correspondent of algebra was based in geometric procedures of cutting and pasting, worked mostly over quantities. So saying that babylonians have somehow anticipated algorithms is a mistake, caused by examining the evidence in the tablets through a contemporary point of view, without fully considering the context in which they were created and where they were applied. Nevertheless, in their numbers as well as in their words, they took a revolutionary step of abstraction of quantities and of categories, in a stabilized system of writing. This step towards abstraction is also a step back from pointing at things towards operating logically with language, looking at symbols and examining them, and developing metalinguistic procedures.

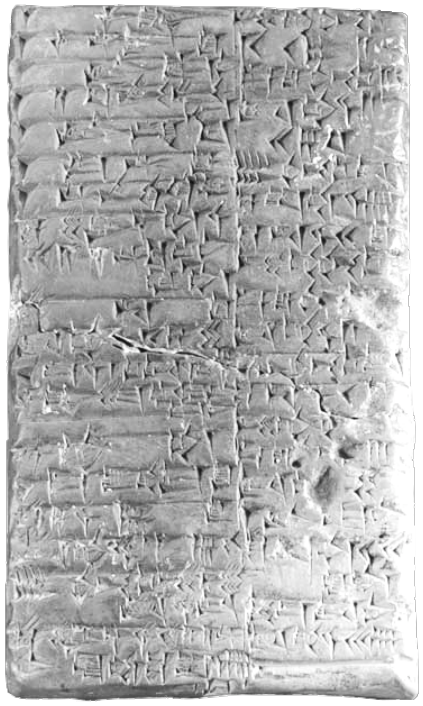


Figure 4.2: Cuneiform script tablet from the Kirkor Minassian collection in the Library of Congress. From Year 6 in the reign from Amar-Suena/Amar-Sin between 2041 and 2040 BC. Source: http://commons.wikimedia.org/wiki/File:Cuneiform_script2.png

So, even when it is organized in structures like lists or tables, data does start already as something discrete, because it is geared towards quantities, it is in the roots of statistics and relates to the discrete character of phonetic writing systems. As we summarized, Gleick (2012) highlights a surprising approximation between logic (or Boolean notation for symbolic logic), calculus and the workings of electrical circuits in Shannon's work in the first experimentations towards digital technologies. Shannon (*apud* GLEICK, 2012), the researcher from Bell Tabs, represents mathematical equations and basic logical operations through series of circuits and relays, where, at every step, the logical paths were defined depending if the circuit was closed or not. It was not a matter of transmitting energy, but of procedurally establishing binary facts, that is, the closed circuit or the open circuit, inside a logical sequence. Following Boole, Shannon had concluded that he needed only two values for these equations: zero and one, for the open and the closed circuit. So, coming back to his mathematical theory of communication (SHANNON, 1948), you get to a maximum economy of signs in a Morse-like system, that is paired up with an electrical circuit, and to which subjective meaning is not central, only procedural logic. So we believe there is yet another underlying detachment that is performed in these experiments: data, or binary values, do not matter for their meaning, that is, for counting or classifying things, but for keeping the logical circuit going and producing more values, more data. They are the engines of the functioning of complex machines.

According to Gleick:

> "Turing was encoding his machines just as Gödel had encoded the language of symbolic logic. This obliterated the distinction between data and instructions: in the end they were all numbers. For every computable number, there must be a correspondent machine number." (p.220)

So data also starts to be measured, and in the most discrete way. The new measure was the binary digits, or *bits*, the smallest possible amount of information. As information is connected to uncertainty, a binary choice, like flipping a coin, would be the indivisible unit. And, to every new bit included, possibilities add up exponentially. As another illustration of the pairing of information and data to the structure of devices, Gleick describes a graph that Shannon once drafted on a piece of paper: an axis for bits storage capacity, ranging from 10^0 to 10^{13}, and along that he organized items such as a digit wheel (used as a desktop adding machine, with 10 digits), a punched card, a typed page, the genetic constitution of men, phono records, a professional journal, the *Encyclopaedia Britannica* and, at last, the Library of Congress (p. 243). Of course, it is an informal note and many of the capacities he assigns miss the mark by several magnitudes. But it is very representative of the transformations taking place, because several different objects (and even institutions) with varied functions and structures are aligned with the human genome and are all quantitatively evaluated by their capacity for data storage.

With digital technologies, data becomes part of a flow, it is a formal resource, something to be managed and shaped, something that also produces forms, information. The difference between data and information is difficult to define nowadays because they are constantly being translated into one another: data can be given shape to be interpreted as information, but it can also be extracted from information and be given new shapes. Both of them travel along networks of devices, but data can be thought of as being less connected to subjectivity and more to formal procedures, while the work of interpreting it leads it towards information.

One of the many contemporary concerns about the use of data is the fact that, once much of our life comes to be described as data, all our knowledge, art and different experiences might be dealt with quantitatively. This would lead to pervasive and technically efficient methods of evaluation, monitoring and surveillance. It might also lead to the development of measurements that, although efficient, do not describe the complexity of our experiences, and might impoverish those same experiences and develop impoverished criteria for the management of social life and for decision making. Or so the argument goes. Marche

provides a good example of this concern in a 2012 essay entitled *Literature Is Not Data*. After describing how Google started scanning books en masse so they could be available for search in Google Books, and explaining some studies on literary works, in which scripts are run in texts by the bulk to identify quantitative patterns, Marche advances that literature is the opposite of data, that what literature is about transcends data and cannot be described by it. He also points out that data precedes written literature, and refers to the same clay tablets described by Gleick (2011), Bottéro (1995) and Roque (2012), with lists of merchandise and their quantities, which he understands as data, while the first written literary work was Gilgamesh, which appeared much later.

Even though Marche's argument touches very significant concerns about how we deal with data and use it to extract or produce information about different subjects, it has, from our point of view, a few problems, that are derived from the idea he has about what is data, and, conversely, what literature can be in relation to it. As Selisker (2012) reminds us while commenting on this essay, almost nothing in the world *is* data, so we should not fear that the researches in the digital humanities, for example, would want to reduce literature to data. Data is a form of description that serves specific goals. It can help in the storage and transmission information, for example. In research, it allows some specific questions about research objects, while in journalism it can be a source for assessing collective or complex events. Most of our experiences, will, of course, present themselves in several dimensions, many of them not describable by data, and neither by literature, for that matter.

In essays like this, in which data is taken as a historical object, there is a dangerous confusion, which we would like to avoid ourselves: of mixing and erasing the difference between data as some sort of recorded hard fact – like a quantity, a measure, a category or a name – and on the other hand as the base for digital technologies. It is easy to blur the limits of what data actually is and start seeing danger in data itself, like if it was some kind transcendent entity or worse, some sort of second nature. As we previously discussed, data is a base for digital technologies in the sense that things are described, stored and transmitted as data in a very broad and detailed manner. It helps formalizing and compatibilizing more and more fields of knowledge, kinds of content and spheres of production, so our society becomes networked. So there is this general impression that everything may be converted into data, because, in a sense, there is actually the practical possibility that every single traceable object has a digital description. But data is, nevertheless, a way of looking at, making and converting things.

So it becomes clear that the difference between data and information is performative, that is, it depends on the ways we approach things and on the interactions with technical objects, on the approximations of what is being said or recorded and its role in giving shape to realities (AUSTIN, 1975). It is not that things were not composed by data in a distant and nostalgic past before computers, where experience would have been fuller, richer and deeper. Previous technologies of information – or technologies of intelligence, if we want to use the broader expression coined by Lévy (1997) – like medieval parchments or even the inside walls of egyptian pyramids with their hieroglyphs, for example, can all be sources of data, if one counts the letters of signs, compares it to other instances, performs measurements and so on, and that does not oblige us to lose sight of the fact that they could be simply read as continuous text and still be used and accessed as a regular parchment or ritual decoration, as instruments of communicating.

Therefore, in the chapters that follow we want to avoid certain pitfalls in the notion of data. First, the idea that data is raw and completely stripped of biases, representing objective reality in direct terms. This is not very precise, in the sense that data, however discrete and atomized it may be, will always carry some structure or the criteria from which it was generated. It is only raw in the sense of demanding treatment in order to be used, precisely because one needs to adapt or clear inherited structures towards other uses.

Second, we should be careful with the idea that data is everywhere: as we have discussed, with computation the meaning of data was added with another layer, that of a procedural resource and formal description of logical operations of software and circuits, that must be shaped and gives shape to information. So saying that data is everywhere, once it is coupled with digital systems, tends to imply that computation is everywhere but it is hard to imagine that computation can stand for the whole of our experiences in communicating, or for all the living meanings of information as described by Capurro. Data does become ubiquitous along with the wide range of digital devices we use, deal with and are exposed to nowadays and it is true that more and more of our personal experience and of social life is being formalized, recorded and transmitted by digital technologies. But, in that sense, information is wider than data, and data itself can be thought of as something older and broader than computing machines. In a sense, data has always been there. We can refer to the description of the workings of the visual cognitive apparatus that we summarized in the first chapter: atomized impressions of shades and colors, or data from the environment, are registered in parallel and aggregated towards useful information. We can also, as we pointed out, count and atomize

anything you like: data is potentially a description to everything, but it does not mean that it necessarily will reduce our experiences and knowledge. Contemporary computers, their software and programming languages perform functions that are much more varied than the logical sequences that are in their deeper base, and the subjective aspects of information are, for better or worse, also present in all the related uses and practices.

This brings us to question a third idea, that data would be like a second nature: as information proliferates in digital devices, data is present in the processes that generate all these forms, so we might get the impression that it becomes a world onto itself that is messy, complex and incomprehensible, but we should always keep in mind that data is produced and that it is in the roots of the visible traces of our exchanges. Data can be found everywhere, but it is not necessarily there, available from the start: it demands an effort to be produced and work for extracting, distilling etc.

One last important observation would be pointing out that, as we already signaled, the categories of data, information and knowledge are not linked to particular devices or technologies of intelligence. It is precisely through this ubiquitous presence that data and information can circulate across devices, being more visible and shareable, which can generate socially shared knowledge, either in a political or a pedagogical sense.

As we have been discussing, the idea of information is developed from intricate exchanges between its subjective or expressive aspects, the objective production of forms and a political and pedagogical function along with many devices that compose the technologies of intelligence. The history of informatics clearly follows a path of progressive *discretization* of information that values formal logic, calculus and data, that is entangled with a correspondent discretization of devices' structures, or at least with the closer engagement of devices with discrete structures.

We would like to advance that this progressive discretization, that has been discussed in the context of the first experiments towards informatics, is part of a process that can be outlined in earlier technologies, historically in the many engagements between cognition, language and devices that inform. We consider it to be an important key to understand the emergence of the contemporary notions of data and information, that is historically conjugated with their visual displays in a similar way in which it has been mapped to transistors and relays by pioneers of computation like Shannon or Turing.

Following this process demands a new differentiation, which is between analog and digital signals: for computational information theory, the first would be continuous, modulated, while the second would be discrete, with no fluid variations of intensity, just fixed degrees. “A rheostat – light dimmer – is analog; a wall switch that snaps on or off, digital” (GLEICK, 2012, p.256). We should observe that according to this criteria, the digital is quite broader than what we call nowadays digital technology. Even though this difference did sustain for a while a separation between cognitive activity and computation, just like with the continuous communication signals and the discrete ones, there is an approximation of both categories through measurement. For example, once you measure brain waves or radio signals in Hertz, you can work out discrete elements from continuous phenomena. This is one of the paths towards the convergence of every media into digital media, and for seeing the human brain and cognition as very efficient information processors. As we have discussed in the first chapter, much of the theory on information visualization works from this standpoint, or at least considering mainly this aspect of cognition.

Deleuze (2003) offers us an interesting perspective on the difference between analog and digital, that takes on a different standpoint, placing both terms as philosophical categories. While analyzing the works of painter Francis Bacon, he draws a discussion that is not exactly about defining clear frontiers between digital and analog, but about the interactions and transformations between analog and digital stances in the cognition that works with images. For him, the digital is already found in the measurements and in counting, it is in the diagrams that are present in a canvas even before it can be filled with brush strokes, it is, above all, in the measuring hand and the recognizing eye that matches schemas, codes to everything that comes along and aligns sequences of interactions, building a determined world, but not necessarily deepening perception and experience. He proposes that the aim of art should be to retrieve an analog language, beyond the digit, “a language of relations, of expressive movements, paralinguistic signs, breaths, screams and so on” (p.113).

Deleuze talks about a diagram, that can be built by the painter in order to make surfaces emerge, as opposed to a code, a preexisting set of divisions of the canvas, an optic schema. The code is “digital in the sense of the finger that counts” (p.104), like the letters in a word. We are submerged in geometry and from our eyes geologic lines emerge even in blank pages, emerge out of chaos, and digitalize things. The world has already been synthesized and it is up to artists to work and sometimes deform these preformatted surfaces to produce analog languages. The painting can build an analogic language in the sense of the hand that gestures

and retrieves continuity and intensities, and integrates senses and surfaces into another landscape. Deleuze assumes that he is pointing to some sort of *reversal* of Platonism, where new transcendents can be bred out of discrete or digitizing perception and devices, and not to the unmaking of codes in order to recede back to some previously existing continuity.

His approach interests us for not being focused on technology, but, at the same time, being very attentive to the effects on the surface of devices and the process of building images, in this case the paintings. He reinserts a fuller sense of cognition and of the processes of subjectivation in the core of the tensions between digital codification and analog languages, while keeping the attention to the surfaces of devices. This allows us to link back devices and technologies of intelligence to processes that lie inside and beyond the borders of specific digital technologies, also retrieving information and communication as social and political practices that move towards building a shared world.

Along the history of the technologies of intelligence, we can identify these many conversions between analog and digital. For example, the history of written language can be seen as a history of discretization of records and signs, that is also influenced by the reinsertion and combination with fluid and continuous modulation. Pictographic systems built sequences of pictures based on direct experience; they cut out and simplify visible objects. On the other hand, the movement towards phonetic writing involved – according to Bottéro (1995) with the example of Mesopotamia between around 3000 and 1750 BCE – a movement of abstraction and discretization, because words that meant something in oral language were broken into specific abstract sounds, each one represented by a letter. In fact, even inside pictographic writing systems we can see this progressive movement of abstraction, where the original meaning of a pictogram may be set aside in favor of assembling a certain sequence of sounds.

The clay tablets of Mesopotamia were marked with cuneiform writing, that displayed more or less atomized signs, and were organized into lists of instructions and tables, in order to reconnect the records that were mainly commercial and administrative. The first works of literary interest, that would emerge much later, will privilege linear and continuous organizations of text, because they are mirroring oral narratives that follow chains of events in flows of sounds.

Carved Greek and Roman writing also display atomized signs. Nevertheless, with the introduction of materials like animal skin in parchments or vegetable fibers and paper as supports for writing and of paint and brushes, we have the emergence of cursive writing, to

which there are evidences even in antiquity. Cursive writing is developed at first for faster recording, and it tends to join words into single complex strokes. The atomized aspect of phonetic writing remains, of course, but is added of a layer of continuous gestures through which the eye flows. For the purposes of readability and standardization, during the Carolingian Empire between the eighth and the ninth centuries, carolingian minuscule writing was established as a standard for communications, and it applied differences between capital and minuscule letters, and all of them were more separated and round, with vertical traces (see figures 3.3, 3.4 and 3.5).

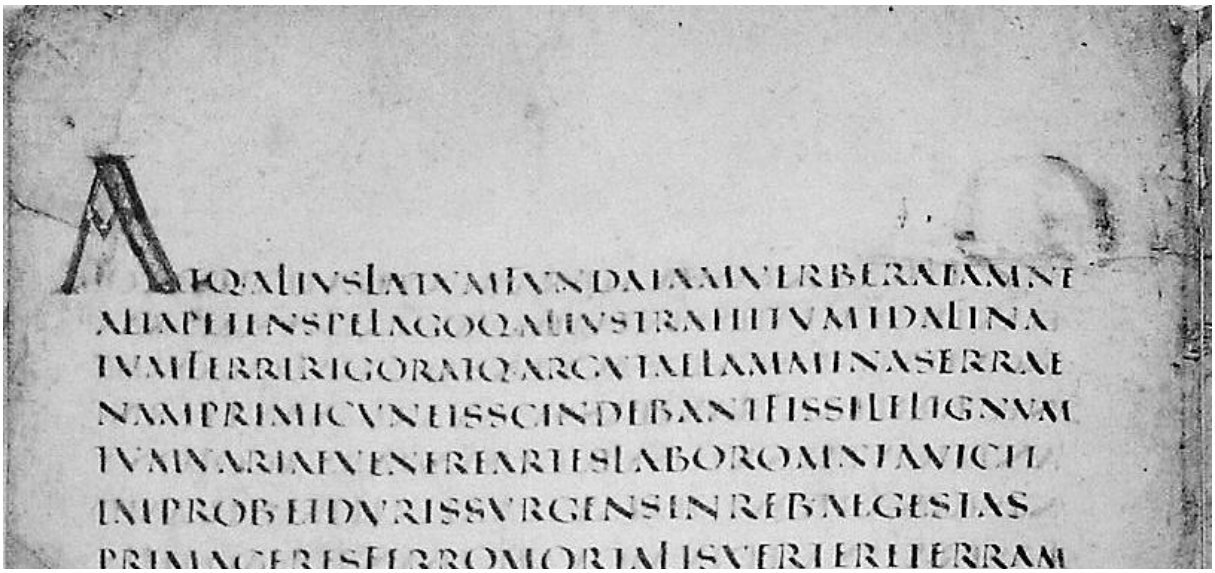


Figure 4.3: Vergilius Augusteus, Georgica141ff, written in capitalis quadrata and in scriptio continua. Source: http://commons.wikimedia.org/wiki/File:Vergilius_Augusteus,_Georgica_141.jpg

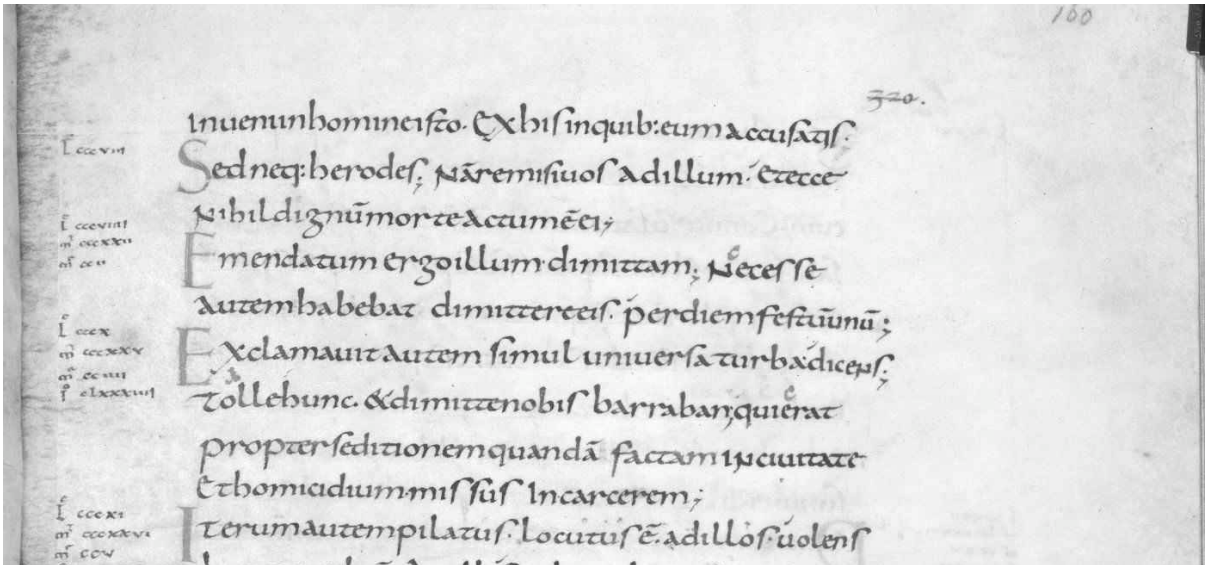


Figure 4.4: Page of text (folio 160v) from a Carolingian Gospel Book (British Library, MS Add. 11848), written in Carolingian minuscule. Text is VulgateLuke 23:15-26. Source: http://commons.wikimedia.org/wiki/File:Minuscule_caroline.jpg

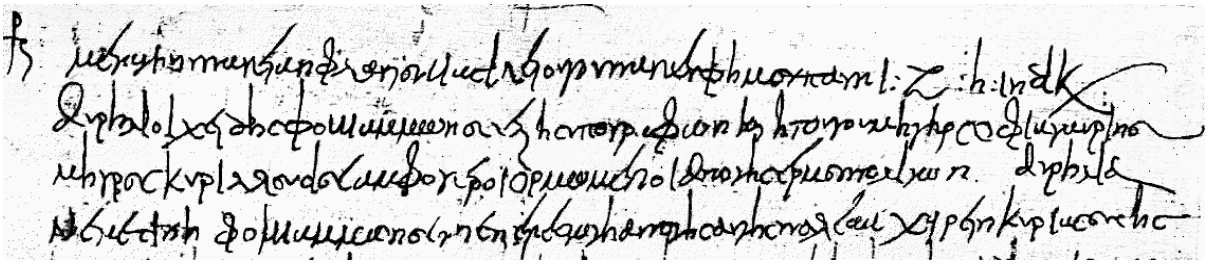


Figure 4.5: Greek manuscript in ancient cursive script, papyrus, dated 545 A.D., Brit. Mus. Pap. 1319 (now British Library Pap. 1319). Source: http://commons.wikimedia.org/wiki/File:Greek_manuscript_cursive_6th_century.png

At the same time, we could trace a history of the development of punctuations, spaces, line breaks and other graphic elements that create separations and at the same time group pieces of text for better reading. Whether in cursive writing or in more formal styles like Classical Greek or Latin, texts were once written without any separation. Around the fifth century BCE, Greeks were using interpuncts to separate words, but it took many centuries for western texts to use full punctuation, around the tenth century. This was associated with the production of many copies of the bible by the medieval scribes, and the efforts of giving visual cues of rhythm and pauses for reading aloud.

But we could also look at the supports in which written text is inscribed and realize that, between the scrolls and the codex there is also a process of fragmentation of surfaces, from one continuous parchment to a series of pages, adding to portability and easing access. Also that with the printed books and more accumulation of pages this will demand page numbering, and maybe indexes: we start building relating tools to navigate larger amounts of content non-linearly, possibly building new continuities.

With the invention of the movable type, we have yet another development towards discretization, but this time it also encompasses the structure of the machine itself, in the types, one for each letter, that should be assembled at each edition, and afterwards reassembled for others. Movable type is seen as one of the great enablers of journalism, and we can see some of its many transformations in newspaper pages (see figures 3.6 and 3.7). Until the end of the nineteenth century, newspapers followed the basic tradition of book printing with single-column pages. Later, most started to print in four or six columns, in order to enable faster reading and to fit more text in simple pages. These columns, nevertheless, displayed continuous text, separated only by the main titles of news pieces. Afterwards, more groupings begin to be drawn, demanding the setting of hierarchical levels of importance

through differentiated graphic treatment, developing complex diagrams and visual codes for the display of the news.

New-York Daily Times.

VOL. I....NO. 1. NEW-YORK, THURSDAY, SEPTEMBER 18, 1851. PRICE ONE CENT.

THE NEWS FROM EUROPE.

ARRIVAL OF THE EUROPEAN MAILS.

AFFAIRS IN ENGLAND.

The Election in France—Arrests, &c.

APPREHENDED DISTURBANCE IN AUSTRIA.

SOUTHERN EUROPE.

Figure 4.6: Cover page of The New York Times. Source: http://timesmachine.nytimes.com/browser

O ESTADO DE S. PAULO

JULIO MESQUITA — RUY MESQUITA

25 DE ABRIL DE 2010 R$ 4,00* ANO 131 Nº 42558 EDIÇÃO DE 20h30 estadao.com.br

DOMINGO

Esportes

Mania nacional
Troca de figurinhas deixou de ser coisa de criança. Pág. E6

Davi x Golias
Santo André pega o favorito Santos. Pág. E1

aliás,
Nova opinião pública

Redes sociais criam uma nova democracia, mostra o sociólogo Massimo Di Felice

Classificados
13.663
ofertas
93 páginas

Novo Porsche Cayenne é mais feroz
Autos

Mercado de imóveis espera seu melhor ano em 3 décadas

Venda excepcional no trimestre indica que a cidade de São Paulo comprará 38 mil unidades novas em 2110

Vereador usa empresa 'de papel' para provar gasto

Figure 4.7: Cover page of Jornal Estado de São Paulo. Source: http://acesso.estadao.com.br/login/?origem=dp

Moving on to more recent examples, with the telegraph people were taking advantage of electric flow to transfer messages using discrete codes but, a few decades on, radio was

broadcasting in waves and was followed by television, only to be more recently converted to digitally-based formats in digital and web television and radio. That takes us to one very particular aspect of contemporary technologies, that is the fact that all analog bases are converging rapidly to digital technologies in a narrow sense, while displaying and allowing for many modulations. Even the devices that did not operate exactly from discrete elements – even though they were definitely measured and calculated –, like the radio or the television, converge into digital technologies. On the other hand, data itself is so atomized that it becomes a flow. It breaks, unbinds traditional structures so they can be fed into the flow: Marche (2012), for example, describes with deep concern the fact that Google, in order to accelerate book digitizing for Google Books built a machine that could cut off book spines and scan loose pages faster. Previous groupings and structures may not withstand the pull of the data current.

What we are trying to demonstrate here is that these variations between discrete and structurally assembled aspects of these technologies bring up issues that have always crossed the borders between language, cognition and devices, either in the sense of instruments used to record something, the object that is invested with information or the artifacts used to transmission and display. Once you think in terms of discretization, it is hard to pull apart what ensues of language systems, for example, and what ensues of the supports or devices for making, displaying, storing and transmitting records. So every translation or agency tends to jump between different classes of objects, and developing a linear history does not seem like the best choice for dealing with the subject.

We understand that developing an approach on these derivations in the constructive interactions between information as communication, data and visual displays will aid us to better understand the role and workings of visualization in heavily mediated societies. We will build this approach as an antiquing work, for slowly gathering a collection of items that may demonstrate these translations along the axis of *discretization* of visual displays towards our contemporary understanding of data, information and visualization. We are using the term discretization in place of digitalization in the sense proposed by Deleuze, in the attempt of avoiding confusion and reinforcing the fact that we will discuss phenomena that is broader than digital technologies specifically. We shall focus on visual inscriptions as traces of the development and the sharing of information. Therefore, our main goal for the remaining of this chapter is to link back information, data, language and visuality before and after the invention of computers and digital technologies in a strict sense. So we are bringing up these

antique objects and intend to place them as landmarks, chosen for representing certain specific issues on the visual representation of data and information. We will discuss their methods in order to better reveal aspects of today's methods, that is, in order to better address these same issues as revealed in today's methods. It is actually a manual craft, sewing work, that would, of course, take much more time and dedication than we can offer now, but we believe even this partial attempt will be very useful for the discussion that follows.

We will organize our itinerary around three main branches: the representation (and building) of time, space and, lastly, the composition of context through visual representation. Each theme will part our path into one lineage, in order to reveal three facets of the same processes of discretization, penetrated by new syntheses and second-level continuities. First, we should gather visual representations of time, like timelines as sequences of events and graphs with time axes, projected on the cartesian plane. The main theme here will be the passage from a continuous and irregular representation of time to time represented in regular and discrete intervals and the emergence of time as monitoring sequences and time series graphics, the issue of sampling and scale. This will take us to a second lineage, organized around the visual representation of space or the organization of territory, with charts and maps. The main issue to be explored here will be the relation between space and territory, and the graphic plane as geometric territory, as the land for spatial measurements. At last, a third lineage would take us to analyze the importance of the idea of context for visual representation, which will lead us to discuss cosmological representations and networks.

Nevertheless, it is important to highlight that each of these three lineages possesses powerful contents that can be related to all the three themes, and also that the themes themselves are mutually defined. There is no experience of time without space and vice-versa and, we would like to argue, there is no experience of time nor space without the perimeter of a context. Therefore, the association between the themes and the lineages and their main components is mostly instrumental. It is mostly guided by the density in which we normally experiment these three themes in each of the components. Our goal and method will not involve reconstituting a historical continuity, but retracing interesting links, this time between visualization and the experience of the three themes we chose.

As a last subchapter, we will talk about lists, tables and grids as three main objects that can be associated with data and the fragmentation of inscriptions, and can also be placed as basic objects for building contemporary visualizations. Along the three themes we will explore,

lists, tables and grids take part in the construction of all the visualization formats we may bring up as examples, and always resurface, sometimes as a *precodifying* base, sometimes as diagrams. Things like lists, tables or grids can be tools for describing things through data, at the same time for producing data and for relating data towards information. We will close this chapter discussing their role in the histories of discretization and, conversely, for the reassemblage of another, secondhand continuity.

4.1. TIME

Time, for Deleuze, cannot be unidimensional: an event, when it happens, is, all at once, present and past, because it happens and in perceived, and because once it happens it is already gone. He sees two forms of approaching the past, related to two flows of differentiation of time: first, past can be seen as preexistence in general, as a memory of temporal regions and layers that coexist, of pasts that are kept in memory. Second, as presents that pass sequentially, forming an accumulation of old presents (1990, p.87-120). Deleuze formulates this image of the two flows of differentiation of time as related to the concept of image-time, that he develops while thinking about the experience of time in cinema. We understand that these ideas encompass experiences broader than cinema and, therefore, become useful as a reference to our discussion about graphic representations of time.

Timelines certainly present this sequential accumulation of past presents, stabilized by all this process of standardization of which we have been tracing some references. Nevertheless, we consider that it would be by far too restrictive to take them only by this second flow of time, of the consecutive presents. Historiographic perspective persists, it is clear, in the flow of pasts that accumulate and in the experience of living while apprehending the fact that one lives. But once the material record of those pasts becomes commonplace, this brings to surface the reconstitution of the past at each access, through reorganizing the virtual reserve of past that is the other flow of time.

Deleuze calls a time crystal this structure in whose center the actual image and the virtual image become indiscernible. This happens in the moment when there is a connection, a crystallizing bond between a present and its contemporary past, that is, between an actual image and its virtual correspondent. Likewise, observing a visual representation of time also mobilizes and crystallizes a core of indiscernibility between the actual data and a reserve of past, in which the timeline takes part, pointing to a bivalence of visual representation. Even if

the straight line itself suggests and is a reference of sequential reading, it is not a straightforward fact that it will be read like that. Moreover, sequential reading, as it often happens, is just in the beginning of the recognition processes. Timelines surround each event in a world of events and crystallize this core of indiscernibility between events with clear boundaries and a virtual reserve of pasts. As it becomes more evident with interactive media, but not exclusively, its most contemporary quality is actually on the rehearsal, in the failed temporal arrangements, in the stalemates… and in the progressive mobilization of these events with other pasts and in the non-sequential reassembling of the extension of time. Depending of the level of detail we choose to use, the evolution of the species can look like it is linear or not: it happens amidst several accidents, rehearsing and encompassing not only bifurcations, but also contaminations and dead ends. Until recently, timelines were mostly the stabilization of temporal course, but what becomes evident in recent times is the incorporation of this cognitive wandering in the structure of the graphic representation of time, so that its presentation it assembled while being accessed. This time order that is multiplied, non-linear, is, finally, an experience of time that overflows from interactive visualizations, in terms of composition and access, but can also be experimented even while building print timelines, with different version for the same timeframe, or the rehearsal of different solutions for an issue regarding time. Rehearsals also go in the interpretation of graphs and timelines, an aspect that is quite useful if we consider turning visualizations into devices for inclusive debate.

Still in the frame of the reflections over the image-time, Deleuze dwells a bit more over this process of rearranging the virtual reserve of pasts at each time we try to build an understanding of an image-time. He relates it to the perceptive experience: we need to place ourselves in the space where things are present, get out of ourselves in order to perceive them. The past is like a general preexistence, that our remembrances suppose and from where they might emerge, while the present is like a tip of the past, indefinitely contracted. Each present can therefore mobilize, by cognitively tracing circles back to the present, different regions of the reserve of the past. The present is the outskirts of the past: it would not pass if it did not carry a certain imprint of the past. The past configures itself as the coexistence of circles more or less contracted of temporal experience. Deleuze compares these circles with regions, deposits or sheets, between which we jump to find this or that remembrance. In these dynamics – where all the sheets coexist in the common limit of their tips of present – would be the foundations of non-chronological time.

Instead of seeing the present as a place for a succession of things, each taking the place of the other, we can focus on the event at hand, and how it is prepared (it is expected), happens and ends, we are taking a point of view that is longitudinal, that is, which crosses a succession of presents towards an in depth view, that distinguishes the event from the space that gives place to it in the present. There is a present in the future, a present of the present and a present in the past, and all of them are tangled in a time that is interior to the event. These are simultaneous tips of present. So, for Deleuze, there are two aspects of the representation of time: the aspects (regions of past) and the accents (tips of present).

So, for Deleuze, present has a plastic characteristic while past has an architectural one. When we place ourselves in this or that sheet of past, we are also conducting a transformation of contractions and distensions, and building a diagram of the stacking of these layers and of points of correspondence. It is like a mental and diagrammatic cartography of time (1990, p.149). In the following pages, we will use this conception of image-time as our main guideline to our antiquing, since it encompasses a rich perspective on the translations of time, on the movements between continuous experience and discrete jumps between specific points of time.

According to Serres (2003), the experience of time has its initial bearing on the organic rhythms and on the path of the sun and other celestial bodies in the sky, as well as the displacement and the relative positioning of the individual himself in the land. It is, therefore, rooted in the body as a reference point for the displacement of other bodies and for the conquest of distances, which demands duration. They are, from the beginning, lines that set up courses, based in embodied experience. Instruments for the counting of time often operate by an impression of those initial lines, like the course of the shadow in sundials, for example. Also, since the most rudimentary machines, by a transfer of these lines to the parts of the devices, like the extension of the ropes in weight systems that measured time in medieval monasteries. So we observe a first displacement of the experience of time that passes, by decal or transference, to useful devices. We believe that, as our experience of time became increasingly connected to many instruments that tell and represent time, to the measurement of time and even to monitoring and its intervals, we come to have different configurations for the relation between time and space. Be it space as distance to be gained by the body or as visual representation. For the purpose of this thesis, we propose a discussion about different visual representations of time and the progressive discretization of its elements towards new

synthesis, which we believe will lead to a deeper understanding of contemporary visualizations and their many transformations.

Rosenberg & Grafton, in their book *Cartographies of Time* (2010), describe in detail the many evolutions of graphic representations of time until the emergence of the timeline, and how this process is related to the development of modern historical time. They advocate that, in general terms, our idea of time always refers to and is involved with the line as an essential geometric form, with lengths and with the displacement along this line, to the extent in which we cannot tell time without the mediation of space. One of the goals of their book is to describe the many experimentations and variations that the representation of time has gone through since antiquity until the modern graphic structure of the timeline was formed, in accordance with the modern the notion of historical time. They begin by discussing the *Chronicle* of Eusebius de Caesarea, a series of books composed on the fourth century comparing four historical chronologies, according to the records of different peoples that were related to the rise of the christian world, and end by discussing contemporary examples and the timeline as some sort of graphic interface for a database, a form of organizing the access to a reserve of information.

According to the authors, the timeline as a visual representation of sequential events is relatively recent, having emerged no more than 250 years ago. Nevertheless, it has roots in other temporal graphic structures and is presented in many different variations nowadays. The first major example that is presented, the *Chronicles* of Eusebius, tries to conciliate different accounts of events of about the same period of time, by aligning each coincident date in the same row. This tabular format allows for an analysis of the differences between each account in terms of what is being recorded for certain epochs and for noticing how different events and people in different parts of the world relate to one another in terms of timeframe. It is a device for putting together different times and it does, after all, demonstrate how different kingdoms were progressively aggregated in the Roman Empire, until the point in time when the message of Jesus Christ could reach all of those peoples.

We should notice that, in order to conduct these comparisons and associations, it was necessary to break up the linearity of the text to present parallel and comparable lists of events, in table cells. It was necessary to make them interchangeable so the sequence in the narrative would be unmistakable, it was necessary to distill the list from the narrative in order

to build a single history from parallel events in different places. So, curiously, in the story presented by Rosemberg & Grafton, the table precedes the line.

They move on, telling us how in the early Middle Ages the most common format for recording events were the annals, listing dates in one column and events on the other. Other than this basic structure, they lacked categories for organizing accounts, making all sorts of entries (or failing to make them) with no consistent criteria. But, towards the Renaissance, chronologies flourished, mixing genealogic trees and historical accounts. For the authors, chronologies at the time were the subject of much dedication and plight, being, in some aspects, more valued than historiography itself, because of the idea that history dealt with narratives, or even poetry, while chronologies dealt with facts, having stronger implications outside the academic study of history. For european christians, it was important to consolidate representations that would match biblical time with the history of european dynasties and even the observations of astronomers. While the authority of many kings was reinforced by their family trees, sometimes stretching until biblical characters, the movement of celestial bodies was seen as the utmost reference of a regular and uniform time. Big historical events were often linked to astronomical ones like eclipses and, after all, Jesus Christ was born under a guiding star. So we believe that this was, after all, a cosmological effort, that aimed to build a harmonious interface between religious, political and scientific concerns.

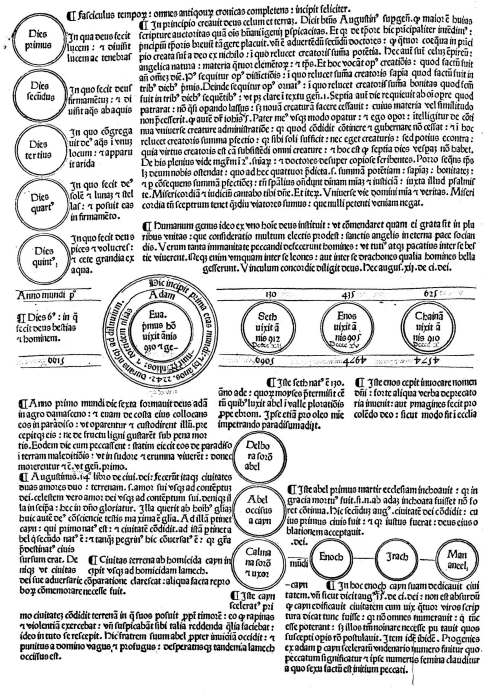


Figure 4.1.1:A page of Werner Rolevinck's fifteenth century *Fasciculus Temporum*, with the horizontal timeline crossing several pages. Source: http://dioscorides.ucm.es/proyecto_digitalizacion/index.php

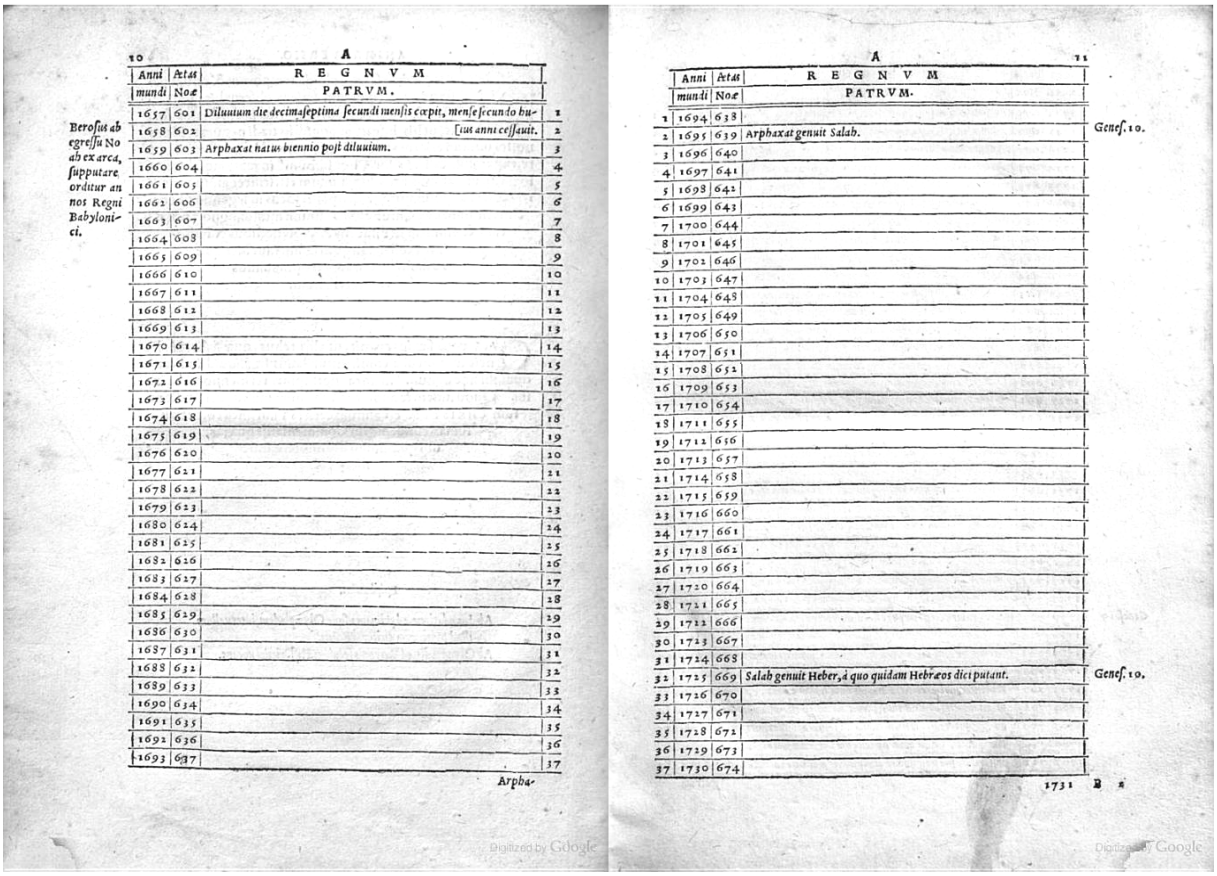


Figure 4.1.2: A spread of Johann Funck's *Chronologia*, from 1545, that used astronomical data to link biblical with classical history. A row for each year. Source: http://books.google.com

And this was a huge challenge: the difficulties of managing detail and scale are apparent in several documents. The solutions found at each situation were quite varied. For example, while some chose to represent royal family lineages along horizontal lines across many pages, with annotations above and below (see figure 4.1.1), others used tables and strictly one row for each year, keeping that regularity even for years where nothing was recorded (see figure 4.1.2). The first emphasizes the rules of different kings, and their beginning and end marked different epochs, but the line was proportional to the details available or brought up for each period, not to its duration. The second emphasizes time regularity and precision, keeping the number of rows tied to the number of years portrayed, but cannot avoid adapting the height of these rows when there are significant events to describe and consequentially more text at a specific entry.

At the eighteenth century, there is a renewed passion for chronologies, associated with the scientific revolution. According to Rosenberg & Grafton, the main advances achieved at that time were innovations in information organization. Many experimentations led to chronologies in the most varied formats, from pocket tables, to scrolls in fancy cases,

sometimes with handles, to discs, sometimes mixing modular structures with ancient visual metaphors. In the *Discus Chronologicus* (see figure 4.1.3), each of the slices stands for a hundred years, from the first to the eighteenth century, and each of the rings represents a kingdom. So, in order to fit more information, the kingdoms that endured the most were organized towards the edges, and the ones that did not last for many centuries or of which there was not much information were organized towards the center. It is, if that is possible, a circular table. It allows for a general view of the records, and this seems to be one of the main concerns for chronologies at the time. Nevertheless, by keeping the strict division between centuries across the circles of all kingdoms, it does not represent visually the beginning and end (also the transitions) between important historical periods, that, of course, did not coincide with the turn of the centuries and differed from kingdom to kingdom.

For the authors, the work of Joseph Priestley would be the most influential for what became the classic structure of the timeline, from the second half of the eighteenth century on. *The Chart of Biography* (see figures 4.1.4 and 4.1.5), for example, combines the years with the names of important kings as an orientation to read the position and the length of the lines representing the lifetime of each historical character. Then, it categorizes the names in the vertical axis, according to the field in which they excelled. It is important to point out that one of the first differences of this chart to the other tables and table variations is that time is actually brought in as a continuous variable: the placement of each life dash and its length do not follow some modular division, but the actual proportional position of the dates of birth and death. So it does actually display the distribution of these lives along the period of time in a much more fluid manner, where the concentrations and vacant areas along the years become self evident. We can, for example, discuss the lack of artists and poets from the eighth to the eleventh century, and the significant raise in those numbers from the fifteenth century on: was it a matter of actually having more artists towards the time in which this chart was made, or was it that the idea of what was art had changed? Many different discussions may arise from it.

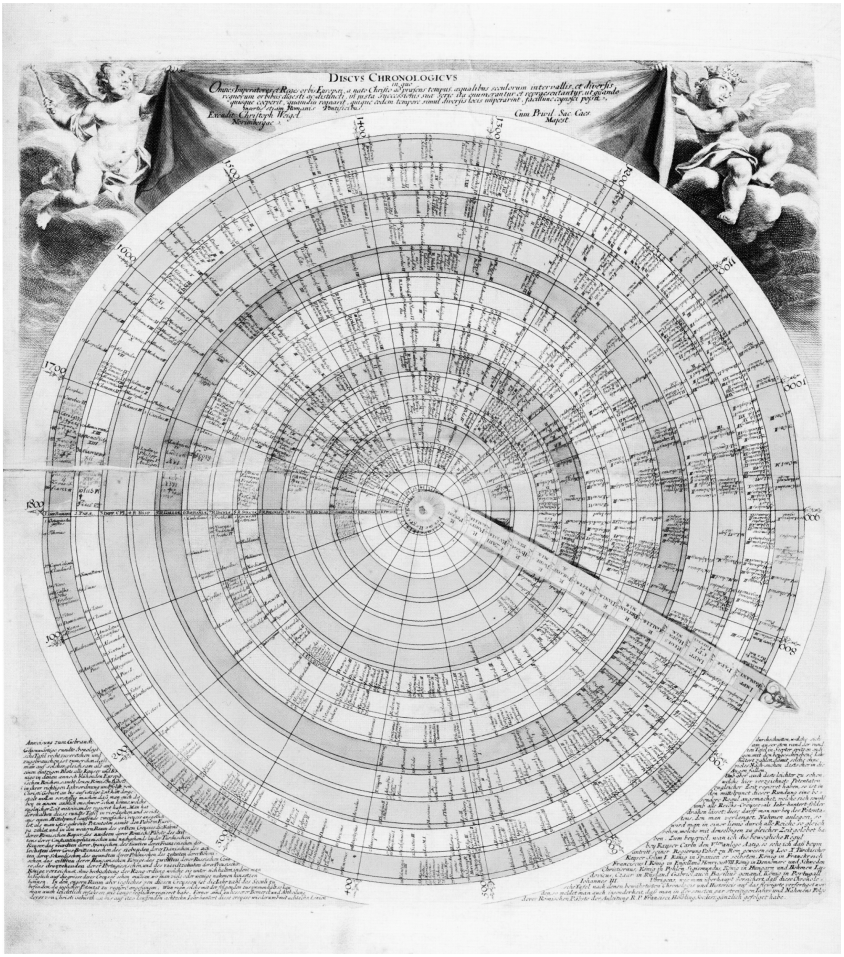


Figure 4.1.3. *Discus Chronologicus*, published in the early 1820s by Christoph Wiegel. Source: https://www.flickr.com/photos/bibliodyssey/4501667579/sizes/o/. *See Annex I, figure A1.1, for color plate.*

Another interesting aspect of this chart is that it is focused on only two kinds of discrete events, people's births and deaths. It strips the temporal progression from preformatted periods and other kinds of events, so the spatial distribution of the ones that are displayed can match their temporal distribution. At the same time, the surface of the chart does not present divisions like tables do, but regular marks or thresholds, orientations that must be followed for positioning the events. With this regular, marked space that has no actual separations, the distribution of the elements in the chart may follow their own values. It is a new kind of chronology, it tells stories with data and is called a chart instead.

We believe that this chart and the rest of Priestley's works represent a very important transition, one that allowed for a closer matching between continuous or categorical data and the space of the display, and that advances towards the atomization of elements. Charts like these would, according to Rosenberg & Grafton, influence the first statistical graphs a few decades later, in the work of Playfair (2005). We should point out that new experiences and

representations of time are emerging, and they demand regular and universally standardized time to be mapped in the regular grid of the geometrical plane, as well as clearer chronological demarcation of events and periods, so they can be defined by discrete values. So we have the combination of this regularized timeline with the geometrical plane at the quantitative time series – in bar, line or diffusion graphs, for example. We believe this will further differentiate the timeline from the discipline of history, and bring it closer to the sequences of data derived from monitoring, that is, of time series of measured samples.

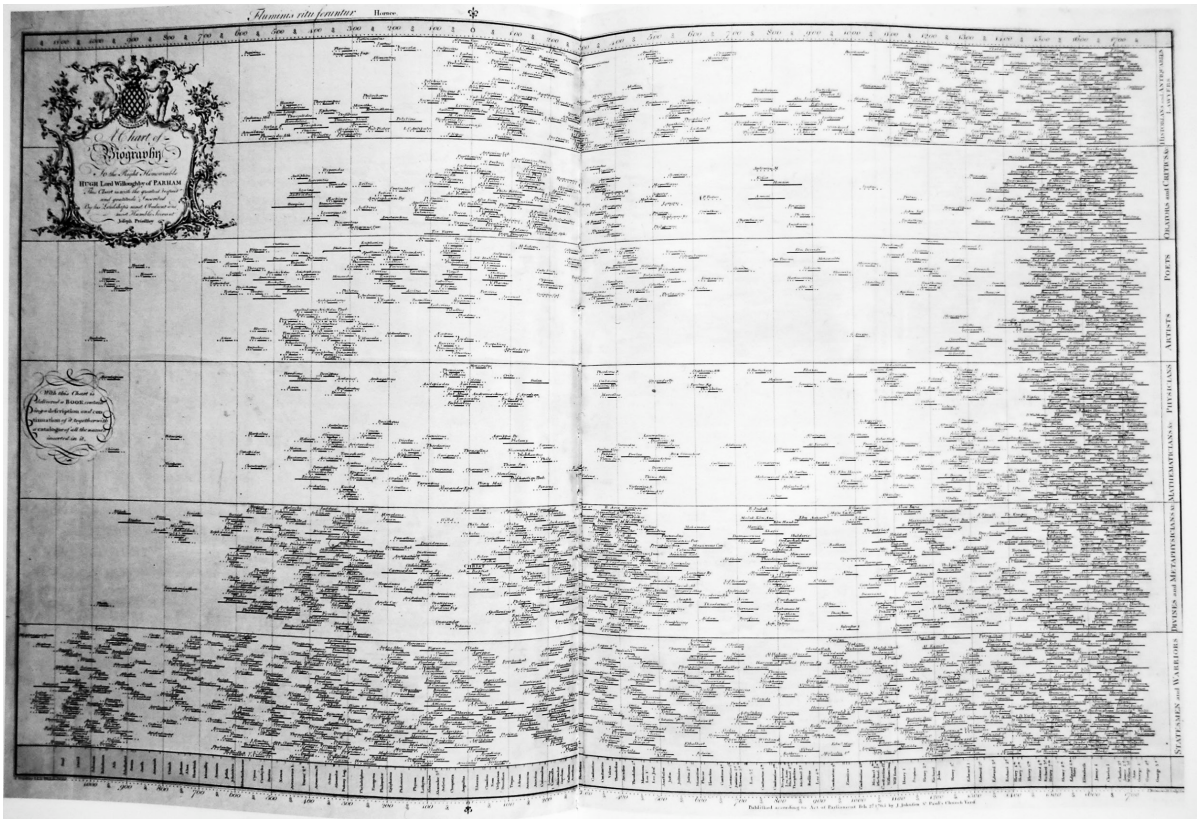


Figure 4.1.4: *A Chart of Biography*, by Joseph Priestley. Source: Rosenberg & Grafton, 2010, p.118-119

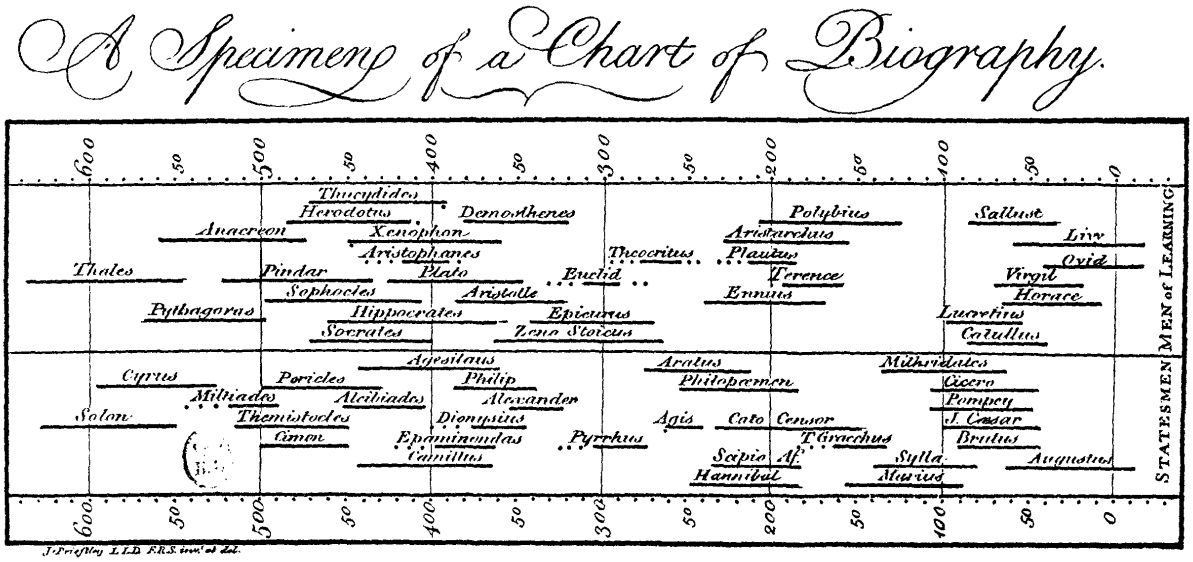


Figure 4.1.5: *A Specimen of a Chart of Biography*, by Joseph Priestley, displaying a fraction of the data displayed of the larger chart. Source: http://commons.wikimedia.org/wiki/File:PriestleyChart.gif

But this movement of developing standards and matching visual representations aligned to them in not recent. Roque (2012) explains that quantification and measure are fundamental for the ideal of knowledge emerging even before Modernity, actually from the Renaissance on. She points out that Descartes advocated that we can only have an understanding of things through quantification and measure, and therefore the idea of an universal knowledge could only be attained through a mathematical base. Apart from that, we would be developing obscure knowledge, and would become more exposed to error. But, of course, this very important movement has two faces. Roque refers to Barbin to explain that, while approaching everything in quantifiable terms, we are working towards turning reality into something quantifiable by separation and triage. These procedures are in the core of what came to be called the Scientific Revolution of the seventeenth century. They are also made visible and collaborate to structure our representations of time.

We relate this triage, this discretization of time to two related developments that were more clearly configured especially towards the second half of the nineteenth century: first, this is crucial for the what we have been calling the management of social life, and for the experience of time from the nineteenth century on. These advances of the eighteenth century regarding the representation of time towards the widely accepted format of the timeline are related with the growing interest in monitoring economic and demographic variables, that is, of checking the evolution of many phenomena along discrete samples taken on regular time

intervals. Second, and this also follows from an interest in monitoring, there is the recording of samples in time by mechanical artifacts. Daston & Galison (2010) described this historical moment when mechanical artifacts like photographic cameras started to be used to produce visual documents with the aim of freeing scientific images from human subjective interpretations and projections, showing direct impressions of nature. They connect this to a specific scientific ethos, the mechanical objectivity. Even though the focus of Rosenberg & Grafton is in the timelines as historiographical artifacts, they also do point to the relation between these and automatic or mechanical forms of recording time in images, like the chronophotography of Etienne-Jules Marey, or the earlier "weather clock" by Christopher Wren, that were emerging as scientific techniques to study movement and monitor variations along time.

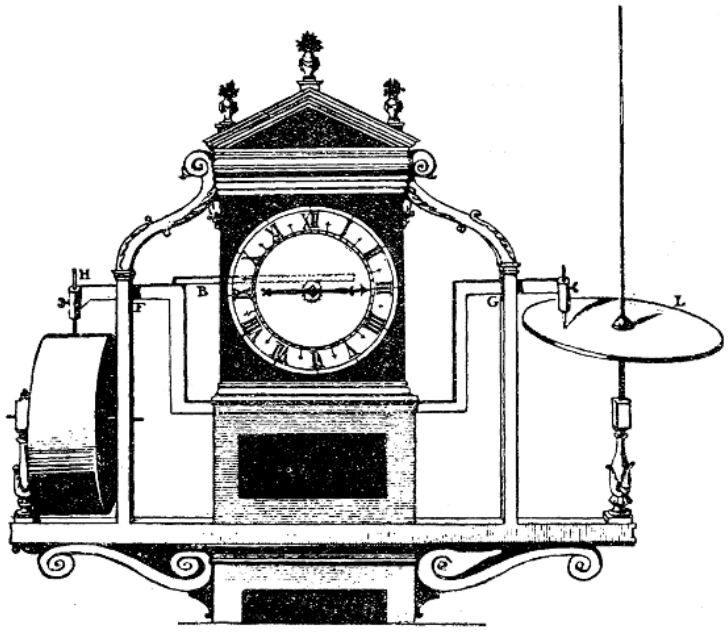


Figure 4.1.6: A contemporary drawing of Wren's "weather clock." Source: http://www.gutenberg.org/files/32482/32482-h/32482-h.htm

The weather clock (see figure 4.1.6) was first sketched by Wren in 1663 and Robert Hooke completed its first working copy in 1679 (MULTHAUF, 1961). It is considered to be the first self-registering meteorological instrument, combining a clock with previous measuring instruments, like a barometer, a thermometer and a rain bucket, to record the changes of the weather along time. This device drew points or lines in a roll of paper and, according to their location, it was translated in terms of different ranges for each variable. Rosenberg & Grafton point out that the creator of the project never expected the raw records of the clock to draw

any interest. Probably for him the marks in the paper were just traces, and the real importance of the artefact was in the finished results, translated into numbers. The great importance ascribed to the device itself was, after all, to provide comparable records across places and times, as with all the early mechanical recording devices.

So by the end of the nineteenth century, when photography started to be more broadly applied to science, the idea of real-time data recording was more or less established, as the weather clock and later variants demonstrate. The first experimentations with chronophotography involved varieties of photographical guns, that would expose the plate at short intervals and document positions of moving objects, animals or persons, their duration and sequence. Many of these informations were not evident to the naked eye, but became very clear once instantaneous snapshots were taken at regular intervals, fragmenting the continuity of movement.

The inventor of the weather clock never wanted to display the resulting image, but the finished, translated data tables. Nevertheless, in raw data instruments this, the image is already abstract. The direct records already represent some sort of schema of the variations recorded, that can be evaluated visually and be used for weather analysis. On the other hand, Photography does not record measurements, but images, so they demand some extra effort in order to be measured and translated into data. So with artifacts adapted from photographical technique, there must be an effort for schematizing the image, so it can collaborate for measurements. This goes for many props used during the photoshoots, like the use of a black background and sometimes black clothes with contrasting white traces so the position of the limbs can be clearly assessed, or to other translations made towards schematic and distilled presentations (see figure 4.1.7).

So, along the development we have been describing, we have the passage from narratives to collections of periods and events, and from that to atomized facts and to periodical markings or samples, which increases to the regularity of the intervals and the preciseness and atomization of each single record. Lines are so blended in devices that the graphic form of the line becomes a device for vision that collaborates for the modern breaching between body and vision. The line in the timeline is not a natural factor, but a trace, like many others, of the relationship of people with technique and technology, of its evolution and mutual constitution.

Figure 4.1.7: Man walking, by Étienne Jules Marey, 1890-91. Source: http://commons.wikimedia.org/wiki/File:Marey-_Man_walking,_1890%E2%80%9391.jpg

Another aspect of this relationship between technology and our experience of time is the displacement of time towards its abstraction, as it also loses its anchoring in the experience of the displacement of the body itself, especially from Modernity on. We see, in different authors like Lévy (1997), Virilio (1993) and Crary (2007), the discussion about how the different speeds to which the body is subject with new systems of transportation and more generally the acceleration of modern life transform time into something elastic, of which we desperately need to keep track, following the lines drawn by the hands of the timepieces and many other paths traced in different devices that will keep on showing us non-elastic time. They keep at reach the reference of measure that is needed for time to be taken as a regular and reliable variable in order to organize different speeds. The experience of time is not anchored in bodily experience anymore, and space, on its turn, is experienced in different ways, depending on the speed available.

With historical accumulation and the efforts for education of the disciplinary societies of the nineteenth century, there emerges a didactic concern: many games and and other trinkets are designed and patented in order to ease the memorization of historical data. Nowadays we got used to delegating our memory to devices of access and storage of information, and tend to value the structural understanding of academic subjects, but at that time, the ability to

memorize dates, names and even classic novels in their entirety was regarded as something necessary for great intelligence. The writer Mark Twain, who also designed games based on historical chronology, posits that dates are valuable assets, but are very hard to memorize; and that this is because they are composed of numbers, that do not have any striking difference in appearance, form no pictures, and so do not profit from the help of the eyes. He trusted pictures to do the trick (*apud* ROSEMBERG & GRAFTON, 2010). So following that line of thinking, many puzzles and board games were created, where historical knowledge was translated into mixing, matching and ordering pieces. Historical time becomes like a repository of discrete events to which only one order will be correct.

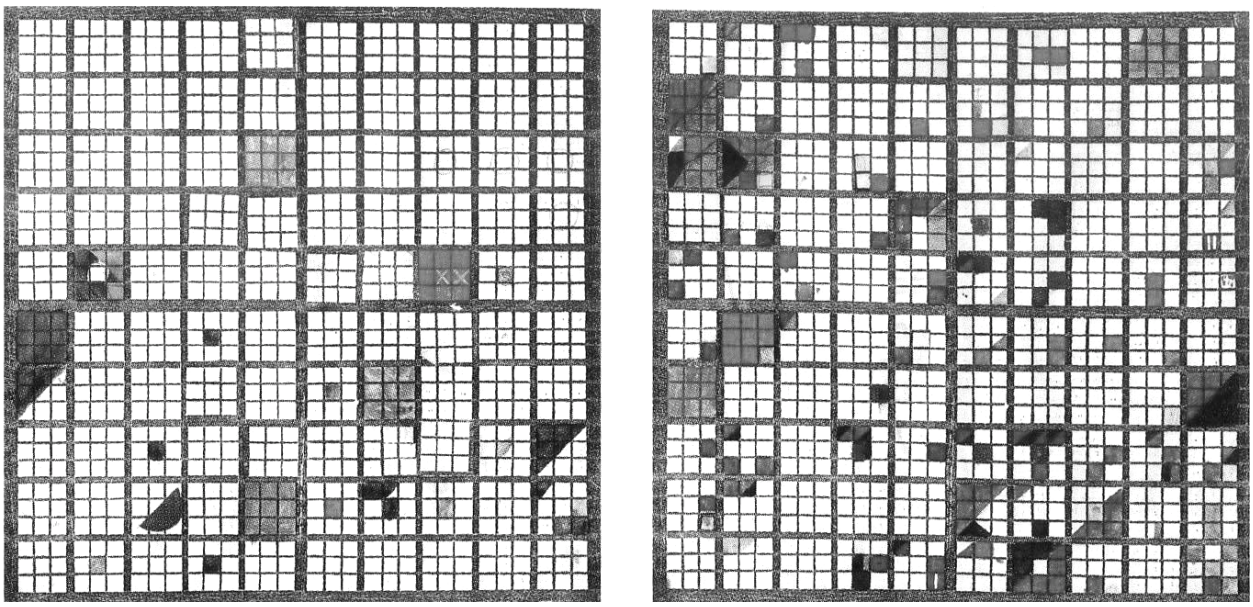


Figure 4.1.8: Elizabeth Palmer Peabody, Polish-American System of Chronology, 1850. Source: Rosemberg; Grafton, 2010, p.205. *See Annex I, figure A1.2, for color plate.*

Rosemberg & Grafton also point out to a very interesting educational device, known as the Polish System (see figure 4.1.8), that was created by Antoni Jażwińsky in the 1820's and over the following decade became very popular at schools in the USA and Europe, especially France. The system consisted of a grid of ten by ten, where each box had three by three subdivisions. The time scale was variable: each box could represent a year, a decade or a century, depending on the timeframe at hand. Each square in the small grids would represent a type of event, to be painted in a specific color, according to the country where it happened. There were also three special symbols, to modify the types of events. This system was used in the teaching of history to improve student's memorization of dates, by mapping them to locations in a grid and creating codified mental images. But, most of all, it was regarded by its

most enthusiastic promoter, the educator Elizabeth Palmer Peabody, as an aid for reading difficult classical texts, as worksheets for thinking through and organizing ideas. Extended cognition at its best: the charts that were preserved “look nothing like one another. Each is the imprint of an individual student’s imagination.” (p.206)

The need of thinking devices only grows once more information circulates as we draw nearer to the twentieth century. The authors point to the fact that the timeline becomes some sort ubiquitous device: “Along with the list and the link, the timeline is one of the central organizing structures of the contemporary user interface” (p.246). Indeed, timelines are one of the main features we have at our disposal for organizing and accessing information nowadays, let alone managing collective activities at work and elsewhere. We would like to add that the graphic representation of time (not exactly historic) has taken up many different shapes, especially since the popularization of digital technologies. Nowadays, it can collaborate for the exploration of richer and multi-levelled organizations of events, like the example of the project Kindred Britain (figure 4.1.9), that brings up a new sense for genealogy: it mixes a network graph of the family links between british historical characters with a multi-levelled timeline, in an interactive interface. Time is also crucial for our access to our own files, by dates of creation, modification etc. Little by little, the past itself starts to be seen as a reserve, waiting to be displayed in different configurations, like memory; and time becomes a variable to be managed, depending on various factors.

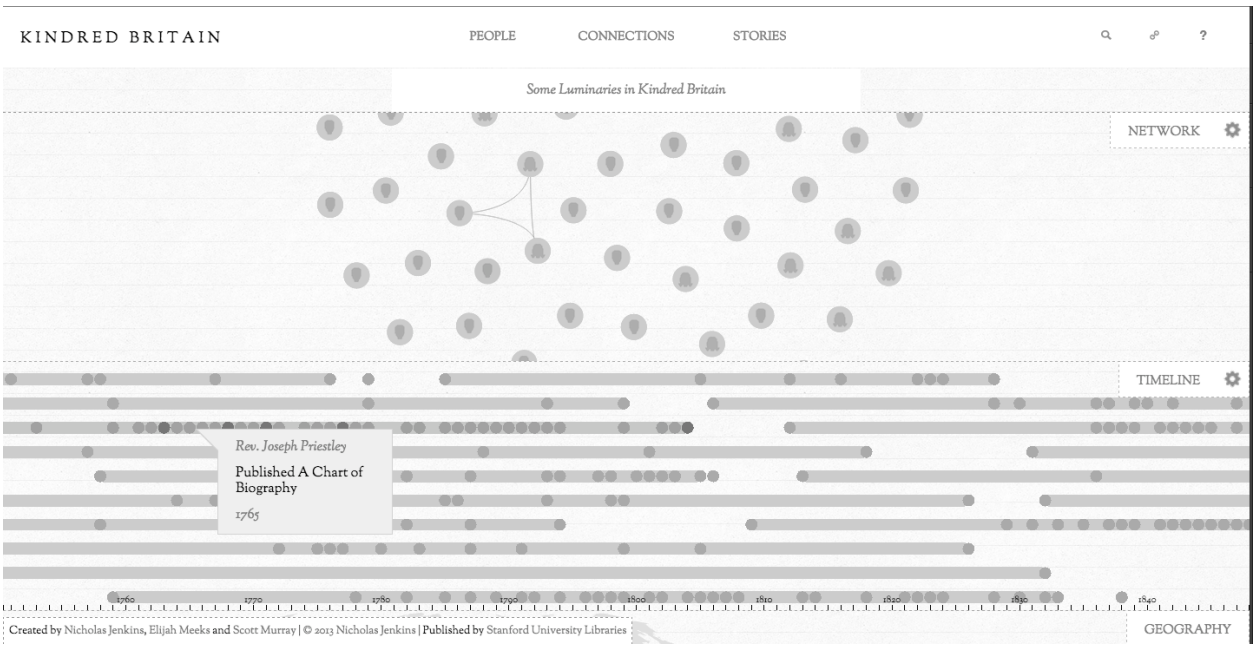


Figure 4.1.9: A view from *Kindred Britain*, by Nicholas Jenkins, 2013. Source: http://kindred.stanford.edu

Speaking from the point of view of the forms of recording knowledge, Lévy (1997), while discussing the role of what he calls technologies of intelligence, defines three poles of the spirit: the pole of primary orality, the pole of writing and the informatic-mediatic pole. Each of them relates to a different way of dealing with knowledge and reveal forms of representing time and expressing it. Even though each pole is more clearly identified with a certain historical period, Lévy explains that they can be pertinent to contexts outside those specific epochs. To put it briefly, the pole of the primary orality is related to societies where oral communication is the predominant form of expressing, maintaining and transferring knowledge, that is, therefore, always embodied in the members of the community. Societies where there is written systems may be nevertheless immersed in orality, if writing is very not relevant for organizing and maintaining knowledge and institutions, and in everyday life. Because of that, to Lévy in a strict sense the establishing of the pole of writing happens, not with the invention of written language, but with the diffusion of the typographic press, that brought writing to common, everyday practices. From that moment on, knowledge is materialized in written records in a broader way, which organize institutions and there is the accumulation of information and knowledge, produced in a progressive manner, aligned with causal historical order. Lastly, there is the mediatic-informatic pole, that is related with real-time media and the permanent modulation and relativity of knowledge.

In terms of the experience of time specifically, Lévy identifies primary orality with circular time, of the natural and ritual cycles, that always return to the present, where there is no accumulation of records. At the pole of writing, there is historical time, more clearly identified with the form of the line, stretched by, in the one hand a past considered archaic where knowledge was unreliable, and in the other hand a project of future and progress. Knowledge and history are accumulated in objects. Because the members of the community do not share the same context of enunciation, there is also a pressure for universality and a corresponding standpoint of the subject, that develops some critical distance. Finally, a the informatic-mediatic pole, time is finally fragmented as a variable, in points and segments, and, because the physical restrictions to the transmission of content weaken, local context is dissolved and is equally weakened the pressure for universality.

When Lévy talks about the representation of time as a circle, he refers to a specific variety of time, which does not accumulate narratives, only resumes and refines them at each cycle, like the narratives in myths that are told over and over. At the height of the accumulation carried out since the beginning of written record, Maeda (2006), in his book about cultivating

simplicity in design, technology and business, proposes the idea that through repetition one may, at each cycle, reach a simpler, more concise formulation. The example is a famous design professor that, at each edition of his summer course, by saying the same thing, each time simpler, “was able to reduce everything that he knew to the concentrated essence of what he wished to convey.” (p.36) In their own way, cycles are alive and well nowadays, as much as the ritual character of teaching, connecting the experience of time with, narratives and knowledge. On the shift to the informatic-mediatic pole, we have a shift from the experience of time as a linear variable for an experience of a multidimensional time, as a network with vertices and many connections.

So this periodization created by Lévy makes it clear that there are other relevant graphic forms for time that are productive for visual representations. We can, for example, bring up some pre-columbian calendars in the Americas, that showed cyclical time weather in circular structures or not. But, we will also have circular representations of time in societies that are mostly on the mediatic-informatic pole, when the cyclical characteristic of time is to be highlighted, depending on the context. The clock is an obvious example, but there are others, like the visualization in figure 4.1.10: it was important to compare the quantities for each similar period, to account for the seasonal nature of the variable, so the circle came up as an adequate solution.

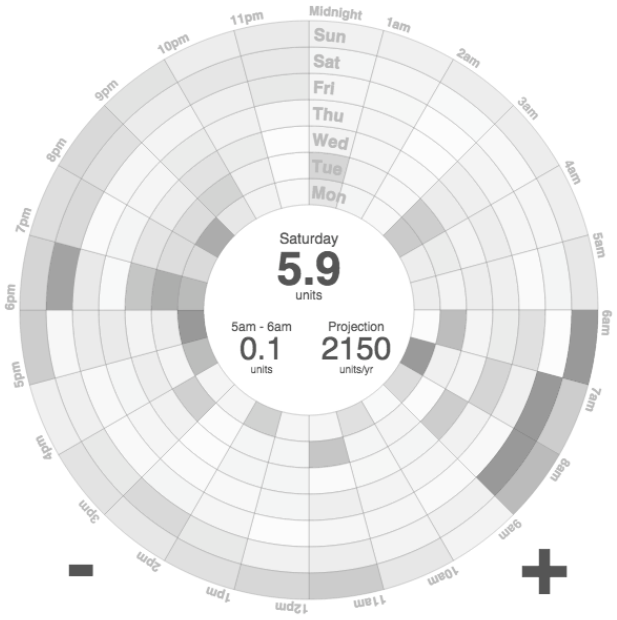


Figure 4.1.10: A visualisation of energy data over the course of half a year, by Peter Cook. Source: http://charts.animateddata.co.uk/energywheel

Apart from the issue of aligning and normalizing time, we should also highlight the issue of the constitution of events and their individualization as elements that can be related and permuted over the regular line of time. There is, as much as with the data within the table cells, a relative reduction of the complexity of the events and of what they mobilize in their respective contexts, so that they can be efficiently represented in a line, aligned, ordered and reordered. Likewise, the structuring of records in tables – classifications, lists, chronologies etc. – is related with the construction of objects of knowledge, in the form of facts, events or data. This idea is related to what Latour (1995) proposes, with the example of the perspective. While Panofsky & Wood (1991), for example, in his classic essay written 1924, highlights the fact that, with perspective, pictorial space has measures that can be transferred to physical space and back, creating an important sense of continuity, Latour thinks the development of perspective and projective drawing is meaningful for the production of knowledge because it flattens and normalizes a plane of representation and individualizes objects. These objects can now be calculated moved around the uniform plane. This allows for more mobility and acceleration for the objects of knowledge, while their integrity is preserved, like mathematical constants. It is like a process of constitution of objects of study, evidences or facts, in a way in which they could be effectively reproduced and verified repetitively in a uniform and flattened context.

Perspective drawing is not, of course, the only visual or graphic strategy for the construction of objects of knowledge: the structure of tables, for example, establishes a grid of relations and also, while isolating the terms into cells, reinforces their factual aspect, and at the same time, it disembodied aspect, as data. Also, graphs in general are also constituted in this double movement: uniformize the plane of representation, that becomes a given field, and align individualized objects, facts to be moved and linked in various ways.

Time series graphs are, from this point of view, a kind of merger between the cartesian plane and statistics that only came to be fully structured by the end of the eighteenth century. Nowadays we have many varieties of this kind of graph, among them bar, dispersion and line graphs. All of them have in common the two cartesian axes (some including also a third axis, for a third dimension to accommodate one more variable), many times with the horizontal representing time. Of course, many of them will not involve a temporal dimension, but this is our entry point for discussing them here.

Even though statistics have been established as a powerful tool for management and analysis of social life, the geometrical plane relates to an older mathematical and geometric tradition. It precedes Descartes himself, and is directly related to the ideals of measure and quantification. In ancient Greece, in aristotelian physics, geometric figures were used to demonstrate theorems starting from abstract problems. These figures were stable, and were devices for a kind of knowledge that did not consider variations of state or movement. This was because, for Aristoteles, movement either was an intrinsic quality of the object, and therefore was cyclic and returned to itself, or was a temporary variation, that would cease once the object returned to the place reserved to it by essence. For Roque (2012), there is a significative change in this approach at the fourteenth century, when some scholars start envisaging the possibility of quantifying the qualities of objects, and propose that they had degrees of intensity that could be studies and, more importantly to us, schematized geometrically. Before that, qualities were classified into pairs of extremes, like wet or dry and, even though they had roots in empirical experience, they could not be measured. An example is the work conducted by Oresme, who understood that a property could vary from a moment to another, or from one place to another. To demonstrate that, he traced graphs where a horizontal line would represent extension (time or space) and the vertical one would represent the intensity of the quality. So, instead of proposing a graphic solution for an algebra problem, as it came to be done later, these graphs present variations according to an extension, using geometrical proportions to demonstrate theorems about the movement and the variation of an object's qualities.

What Oresme did, even though his work was in general terms in the bounds of aristotelian physics, where movement is also a quality or a temporary variation from natural order, was parting extension (either time or space) and intensity (chosen qualities) into two different but related streams. By this division he was opening the possibility of turning them into quantifiable variables, that is, to make further divisions and count and relate indefinitely. So these figures are one of the basis for today's graphs, where the basic variables will be plotted as extensions in the graphic space.

Galileu is going to deepen the use of diagrams to represent the variation of quantities that happens in movement, mapped out in geometrical forms. According to Roque (2012), for him, any qualities that could not be quantified did not have relevance for scientific knowledge and should be disregarded. Nevertheless, his diagrams will still be disconnected from algebra, and

will not serve as an instrument for algebraic demonstration, which would be the fundamental innovation of the cartesian plane.

So the great merit of Descartes, in the context of our discussion, was instrumenting a new art of invention, that is, aiming at theoretical devices that would serve mathematics in the same way as technical objects served social life in general (ROQUE, 2012, p.318). So the demonstrations should generate knowledge about the nature of the problem and propose paths towards its solution, instead of clarifying and perfecting certainties. Also, they should propose some sort of experimental context for the resolution of practical problems. Therefore relating a geometric curve of the possible solutions for an equation and the trajectory of a projectile, for example: demonstrations close in on simulations and project back their lines and quantities into our understanding of the landscape. Also, it will be the basis for today's computer vectorial graphics.

In the graphs with coordinate axis nowadays, one of the lines, as a basic and unidimensional geometric form, will bring back regular extension to time, or even to space, while the other axis will present densities for each corresponding point. As Deleuze puts it, movement can only subordinate time and turn it into a number that indirectly measures it if it is somehow normalized (1990, p.50). Returning to our main subject in this chapter, movement, in our case, is the enchainment of events over the normalized line, so time shows up as a number of this movement, in the sense of being its measure.

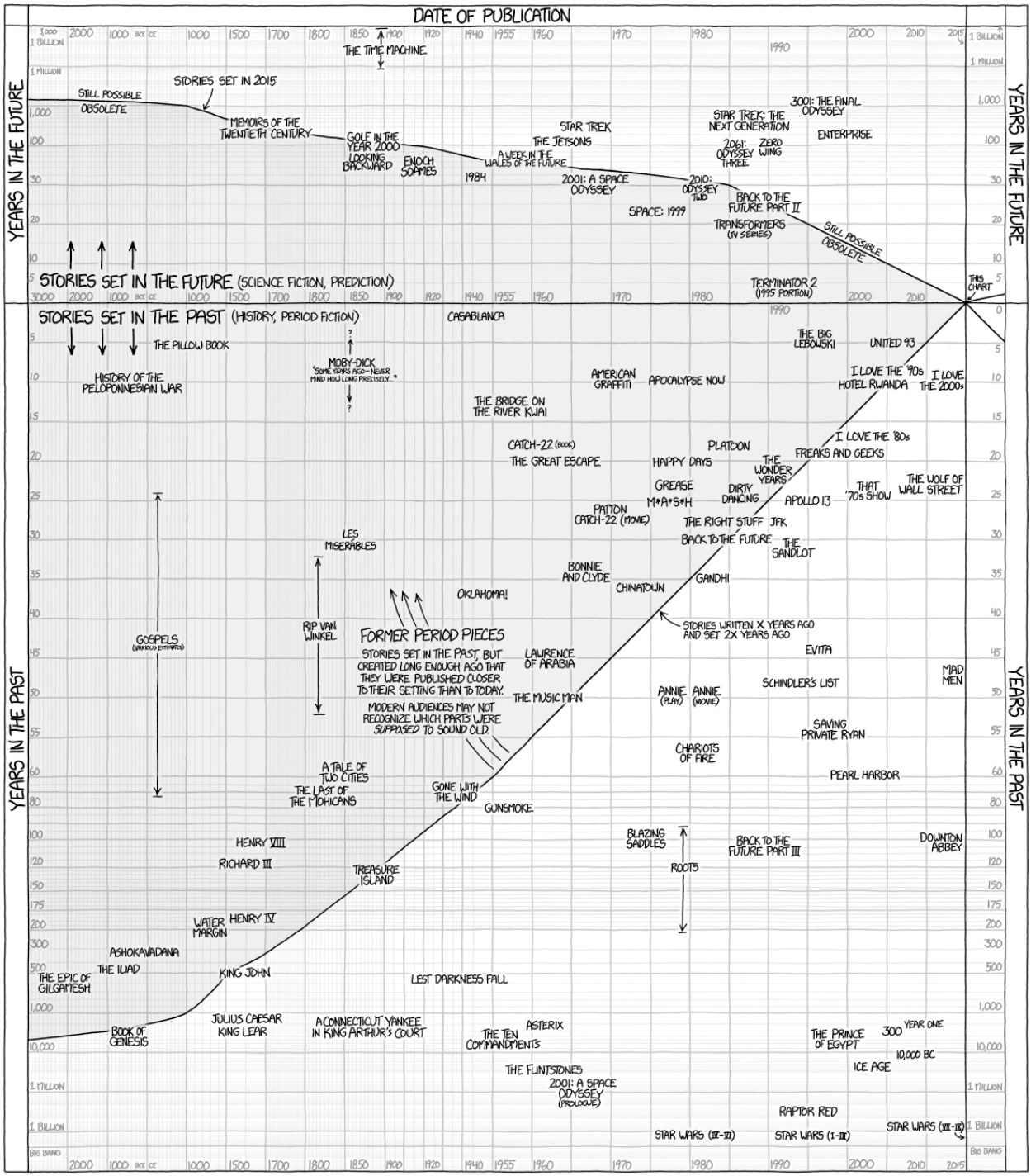


Figure 4.1.11: *Stories of the Past and Future* by Randall Mundoe, 2015. Source: http://xkcd.com/1491

If historically the basic line has been maintained as the central axis for the visual representation of time, timelines have assumed many different presentations, becoming at times multileveled, spiraled, cut into connected sections etc. Without dissolving the line as a basic resource for conducting transformations and linking, timelines are no longer necessarily centered around a central axis, but oriented by many. Network maps, for example, can display

sequences of simultaneous events: in social media analysis, links may display sequences of system requisitions like retweets, in public health research, sequences of contagion. There may also be layers of superimposed times, like the graph crafted by Randall Munroe in his online comic strip, were fictional time and historical time, past and future, meet in the same plane (see figure 4.1.11).

As representations of time are connected or generated by systems of information, also being structured in real time, the access to them becomes a negotiation of how to map time, and we are closer to the sort of experience of time described by Deleuze, as something that may be architectural or have a certain plasticity. Visualizations start, finally, to link navigation with a topological experience of the reserve of past that complement our own memory. Of the many approaches one can have towards a timeline, we start to refine a point of view where the presentation itself in an issue, it is part of the discussion of the events at hand. Instead of experiencing spatialized time, we come use timelines as devices to experience temporalized space, where configurations and transformed and are problematized. This is especially true for some forms of interactive visualization, but is also part of an approach that can change how we use traditional ones.

4.2. SPACE

As we observed in the previous sub-chapter, representing space is closely connected to representing time, and the relationship between both dimensions of experience is always rearranged as we represent them. In the pages that follow, we will discuss some historical variations of the visual representation of space, especially cartography, and its connection to contemporary visualizations. Classic cartographies may emphasize the description of physical geographic features and political frontiers, or these may be used as a basis for visually displaying the geographic distribution of other information, like in statistical maps. At a first glance, this second configuration seems more connected to modern forms of management of social life and to our work, that involves discussing the elaborations we build from data. But history shows that both perspectives are quite ancient in mapmaking and inform much of our experience with space in contemporary visualizations.

According to Harley & Woodward (1987, 1992, 1994), who organized a collection of books titled *The History of Cartography*, even though maps or ancient variations of maps may be said to exist since prehistoric times, the term “cartography” or even the term “map” still lacks

some clarification. Studies of the history of cartography have been either accepting the terms uncritically or settling for very restricted definitions, understanding cartography as a depiction of the surface of the earth, or a part of it, therefore tied to a narrow conception of geography. They also explain that, despite the important role of maps in ancient societies, there are no specific terms to designate them in ancient languages. The roots for the english term "map" are in the latin word *mappa*, meaning a cloth. While many languages share this same root, the correspondent term for map in others is derived from yet another latin root, *carta*, which meant any formal document. These same terms or variations of these are still in use, designating other objects than maps. We could also point out the use of the term "chart" in the English language for maps for sea navigation and tables at the same time. So even the question of what is a map brings up many issues.

The authors settle for the following definition: "Maps are graphic representations that facilitate a spatial understanding of things, concepts, conditions, processes, or events in the human world." (p.xvi) They point out that maps are not restricted to depicting physical space or land, and that elaborate cartographies have been composed to depict worlds of dreams and speculations, and these should be part of the history of cartography, as well as celestial cartography and maps of imagined cosmographies. They explain that this definition matches both their concerns "with maps as artifacts and with the way maps store, communicate, and promote spatial understanding" (p.xvi), and also avoids the understanding of maps exclusively from the point of view of a historical-literate experience and from the criteria of specific methods and procedures, like latitude and longitude or those connected to systems of numeration or metrology.

On the other hand, it is important to notice that the spatial understanding of all the things they list in this definition does not seem to be necessarily restricted to geographies, nor to any previous notion of space that would be depicted in the map. As we have seen in the previous sub-chapter, the visual understanding of events may involve a graphic, spatialized representation of time. So of course we can argue that timelines produce a geography of time – hence the title of the book *Cartographies of Time*, by Rosenberg & Grafton (2010) – , but then the definition offered by Harley & Woodward would fit for any and all visualizations. Nevertheless, when they move towards defining cartography itself and the scope of their books on the history of cartography, they say that it should be necessarily connected to:

> "the historically unique mental ability of map-using peoples to store, articulate, and communicate concepts and facts that have a spatial dimension; and (...), since

> cartography is nothing if not a perspective on the world, a general history of cartography ought to lay the foundations, at the very least, for a world view and its own growth." (p.xviii)

This necessary spatial dimension they identify in cartography is, therefore, to be considered as something anterior to mapping, or at least that relates to other spaces outside the map, like a spacial distribution or organization of things, either measured, conceptualized or imagined, that should be depicted somehow. In this chapter we use maps to discuss this relation, between the spatial dimension of experience and its constitutive depiction.

Although there are many scattered evidences of prehistoric representations of space, like drawings of valleys with contours of hills and villages and lines of rivers, or the arrangement of houses or even of stars in the sky, it is hard to tell at which point these landscapes could serve similar purposes of maps, so the origin of maps is quite nebulous. According to Smith (in HALEY & WOODWARD, 1987, p.45-49), the study of this material has been insufficient, mostly because of scholars' lack of appreciation for the relations between these quasi-maps and the cognitive development of early civilizations and the needs they faced, that were different from those of modern times. For example, Lewis (idem, p.50-53) posits that, contrary to modern scientific awareness, where the focus is on order and regularity, where maps should display with the most precision what is known and progressively amplify its range and detail, prehistoric people had to maintain a constant state of alertness for the unanticipated and unexpected. Therefore, prehistoric maps focused on irregularities and on uncertainties rather certainties. On the other hand, there are also ancient cosmological representations that seeked to represent spatial dimensions, even if it was a space disconnected from wayfinding.

So we have this evolution in the cartographic field, from a moment in time when there was not a clear distinction between on the one hand ritual and teleological representations and on the other hand instrumental maps, for aiding hunting or mobility; to the many different representations of space we have nowadays, like astronomical maps, world maps, maps with physical information on the land's surface, statistical maps, along with cosmological representations, maps of fictional and onyrical lands and many other variations. Each of these will circulate in different environments and have different criteria for attributed value and meaning.

There is a long path of transformations from ancient maps that combined empirical knowledge with symbolic aspects; and the current state of *cartographic specialization* that

produces many different forms, but elevates the measured decals produced by modern cartography as the pinnacle of cartographic works in general. The maps and map variations in ancient Greece display a strong example of this change of perspective: even though the Greek civilization can be said to have existed since 21.000 BC, the first evidences of map-making date from about the sixth century BC. According to Aujac *et alii* (in HALEY & WOODWARD, 1987, p.130-147) the earliest maps were influenced by Homer, who is to be considered the first geographer in history. Even though he did not produce any maps – any maps that accompany Odyssey and Iliad were produced afterwards, from interpretations of his writings –, his rich geographic descriptions influenced subsequent map-making in Greece. They posit that the contribution of Greek cartography lie more in the speculative and theoretical realm than on the practical realm: terrestrial mappings lacked precise direct observation and astronomical projections relied more in abstract geometry than in methodical measurements. Even though from the Hellenistic period on geography was regarded specifically as a description of the earth, for the authors it is clear that for the Greek mapmaking also included a concern with depictions of the universe as a whole.

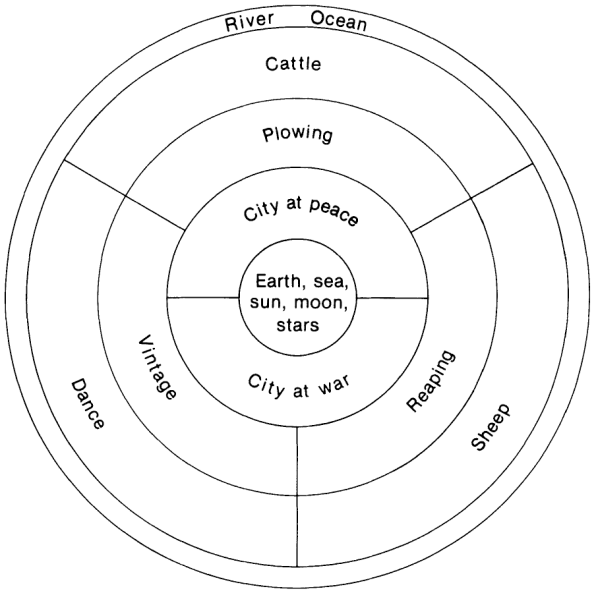


Figure 4.2.1. Reconstruction of the shield of Achilles from Homer's Iliad. HALEY & WOODWARD, 1987, P.131. *Apud* WILLCOCK, Malcolm M. *A Companion to the Iliad*. Chicago: University of Chicago Press, 1976, p.210.

From the start, they draw attention to Achilles' mythic shield (see figure 4.2.1), described as a round piece composed by concentric pieces: in the center, depictions of the sun, moon and star constellations summarize the astronomical knowledge they had at the time, and around it, two cities, one in peace, the other at war, and further out rural scenes, and everything

embraced by a circular sea. According to Aujac *et alii*, “Even in this poetic form we can glimpse the use of a map, almost as a heuristic device, to bring some order into concept and observation and to codify the early Greeks' reflections on the nature and constitution of their world.” (p.132)

Nowadays it may seem mysterious to us that this schema places the sky as the smallest part in the center, because we are used to imagine the sky as encompassing everything, not the sea. We believe this shield is part of a world vision that is anterior to the bird’s eye standard we have for mapping, either in regular, measured maps and in cosmological representations. We would like to advance that it depicts the world through the vision of someone who is planted on the soil, on earth, and, at the same time, dominates all the habitable world until its limits in the vast sea and up ahead, sees the sky. Examples of similar points of view, that are not evident for us nowadays, can be found, for example, in religious paintings that portray the heavens in medieval churches and in a contemporary and less biblical perspective, in the pictures of Randy Scott Slavin through 360 degree fisheye lenses (see figures 4.2.2 and 4.2.3).

Figure 4.2.2. Ceiling of Duomo di Parma, "Assunzione della Vergine", by Antonio da Correggio. Source: http://commons.wikimedia.org/wiki/File:Cupola_Duomo_Parma_Correggio.jpg

Figure 4.2.3. Photography with 360° fisheye lens, by Randy Scott Slavin / Rex Feature. Source: http://www.dailymail.co.uk/news/article-2131638/Who-needs-Instagram-Photographer-uses-traditional-fisheye-technique-capture-extraordinary-landscapes-America.html *See Annex I, figure A1.3, for color plate.*

Even afterwards, in the maps proposed or produced by Anaximander and Hecataeus in the sixth century BC, this round shape has been the most widely adopted, with a flat land and one single ocean embracing all of it. Hecataeus drew a fairly faithful description of the world known to Greeks at the time, claiming to be an improvement of Anaximander's, to which there are only references nowadays. It was already composed from a bird's eye point of view, and divided the world in two continents, Europe and Asia, this later including Libya, at the time a general name for the African continent. Even though his map and its variations depicted the coastlines of the Mediterranean shore with relative precision, the units associated to it were not linear, but associated with time: days of sailing on the sea and days of marching on dry land. Its round shape was soon to be questioned, not to say ridiculed, and caution was warranted because of the non-regular measures it displayed. Nevertheless, these circuits of the earth continued to be used much longer, even when most already knew the shape of the continents was much broader in every direction, and that the proportions were misleading.

It was Herodotus, in the fifth century BC, who systematically challenged this general configuration, advancing that mapmaking should be based on experience rather than geometry. He argued that there was at the time no clear knowledge that there was water around the Asian continent, for example, and that there was no clear motivation or criteria for tracing the frontiers between the continents and advocated that lines should be traced according to positions of geographic features. Also, he saw no sense in drawing maps that encompassed the whole earth, since it had not been fully explored. Maybe for the fact that Herodotus never drew graphic solutions that supported his argument, he is not considered to be a geographer. But we can say that he, nevertheless, collaborated to a movement towards the combination of theoretical knowledge and direct observation: he "seems to have urged a return to empirical cartography founded on exploration and travel. Theory, in his view, should give way to experience." (p. 136). Yet, it was Aristotle, in the fourth century BC, who established that the Earth was spherical, combining observation and geometric theory: he observed that, for example, the lunar eclipse was always circular and that ships seem to sink when they reach the horizon, and that some stars can be seen only from certain parts of the Earth. So he mapped different areas or slices of the globe and developed geometrical demonstrations of seasons, winds and astronomical events.

From that point on, Aujac *et alii* relate the progressive establishment of an empirical cartography in Greece with new political, military and cultural developments in the society as a whole, especially with the role cartography starts to have for the growing educational structures. Also, with more exploration and conquests, the Greek world expanded and this demanded an organized body of knowledge about it. For sure, there is evidence that some varieties of maps were used before that time, at least in the Greek polis, in a microscale: maps for mines, for travelling, for depicting the city. Maps for everyday life that had a much more practical sense. Nevertheless, it is with this expansion of the Greek world that the first systems of orientation, coordinates and latitudes start being defined by Pytheas, expanding the work of Aristotle, which led to more precision in macro geographic projections.

Eratosthenes took these developments to an experimental level, when he measured the circumference of the Earth. He argued that accurate maps depended on linear measurements and developed a system of meridians for charting with more precision. But it was Ptolemy who developed a method of perspective projection laying out a grid of coordinate parallels of latitude and meridians of longitude over the globe. This uniformized scale and measures and proposed an equivalence between cartographic space and physical space. Lines on the map

ceased to be organized by surface measures, incorporating projections of this grid. Dilke (in HALEY & WOODWARD, 1987, p.177-200) also highlights that the most crucial legacy of Ptolemy to the development of cartography was the way in which he codified instructions to how maps should be drawn. This is related to this discretization of the surface of the globe in grid units:

> "To make a simple analogy to modern cartographic data bases, we might say that Ptolemy transmitted his cartographic knowledge in digital rather than graphic form, leaving his successors to recreate the images he so clearly envisaged as the end product of the mapping process." (p.180)

The author also discusses Ptolemy's cartographic tables: lists of locations and their coordinates: we can see that the astronomer and geographer created a discrete system that allowed for many transformations between structures. From the models of the globe to a grid of coordinates covering its surface, to data tables and back to surfaces in the maps (see figure 4.2.4).

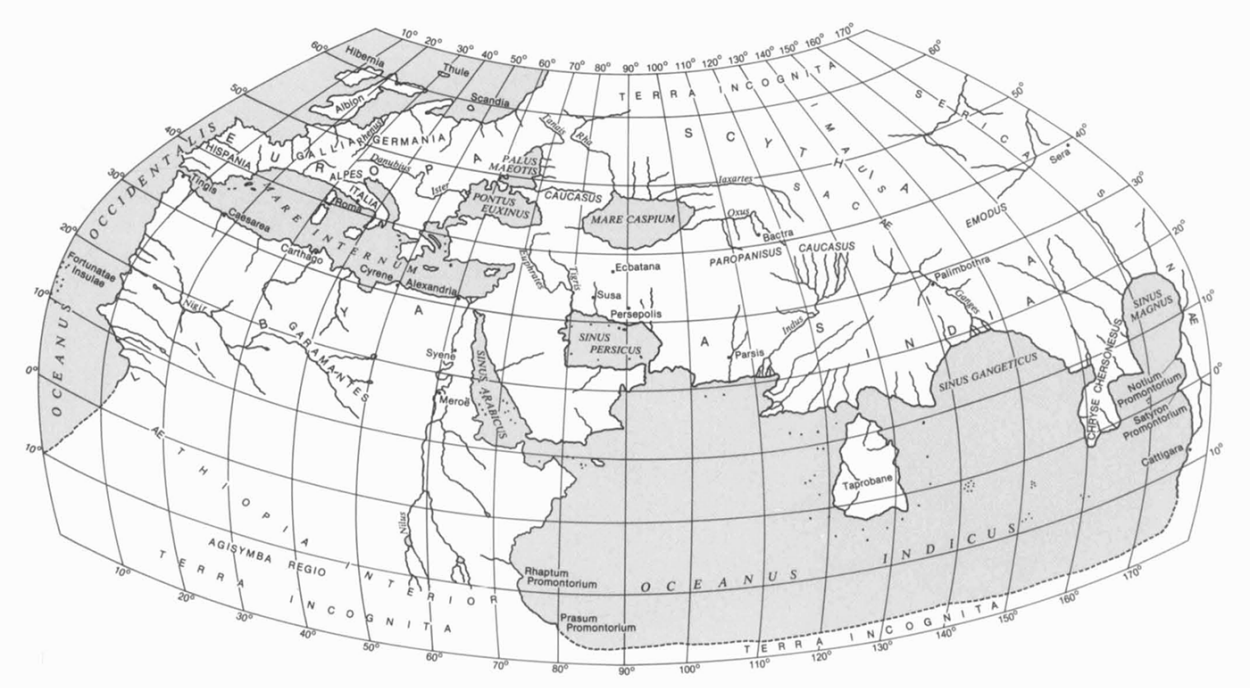


Figure 4.2.4. A Reconstruction of the World of Claudius Ptolemy. Source, Dilke, in Haley & Woodward, 1987, p.185. *Apud* Edward Herbert Bunbury, A History of Ancient Geography among the Greeks and Romans from the Earliest Ages till the Fall of the Roman Empire, 2d ed., 2 vols., 1883, p. 578.

In parallel to these developments in Greece, there were other cartographic techniques being developed in China that display yet another interesting aspect of the development of cartography towards modern uses. According to Yee (in HARLEY & WOODWARD, 1994,

p. 96-127), we could say that chinese cartography went through a different development than greek or western cartography: they worked on complex systems of representation for economic data, that was integrated with measure, and had also different criteria for representing topographical information. An example of this first difference are the cartographic works of the Han dynasty, a set of seven maps in wooden blocks, dating from around the fourth century. These display overlapping spreads of land, and indications of administrative counties, rivers and roads, along with sites for gathering different types of timber, and sometimes distances between them. They account for an area of approximately 107 by 68 Km, and can be considered the oldest economic maps, crafted long before Strabo's maps, their first western equivalents in this sense. There are also three maps from the Han dynasty, found to be a bit more recent and complex, for covering a larger area and displaying more information, including figures on populations and military sites, along with distances between landmarks. These annotations could be quite elaborated and codified, as the reconstruction in figure 4.2.5 shows.

We can identify a higher integration and dependency between textual information or figures and the spatial depiction of the land, which was fruit of much plight in developing cartographic language, but, on the other hand, showed a lack of precision in the reproduction of physical features and in the scale. The use of regular grids for mapping would only come with cartographer Pei Xiu, in the third century AD. According to Yee, Pei Xiu went to lengths to develop drawing methods that would narrow the gap between observation and depiction: he observed, for example, that mountains and hills were depicted all in a uniform manner, with no regard for their varying heights or base widths and other physical features. As Chinese maps at the time of Pei Xiu, due to aesthetic traditions, had an oblique point of view (and not really perpendicular, like the Greek ones or like most modern ones), he developed a complex system of geometric projections and equivalences to depict these varying proportions. One of his major contributions was managing to display these topographical elevations with more accuracy. By the beginning of the seventh century, the Emperor Yang commissioned officials throughout the empire to document the customs, products, demographics of all provinces, that were to be used in the development of many contiguous statistical/economic maps. Those were maps of peoples that reinforced the human aspect of geography in ancient times through cartographic systems.

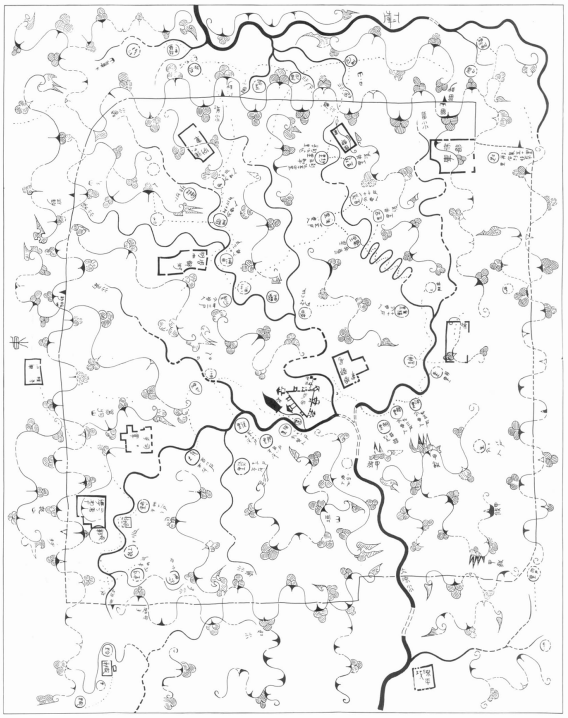


Figure 4.2.5. Reconstructed Garrison Map from Mawangdui, from the second century BC. Source: Yee in Haley & Woodward, 1994, p.45. Apud Gu ditu lunwenji (Essays on ancient maps). Beijing: Wenwu Chubanshe, 1977.

These point to the assumption that maps in China were, since the first cartographical attempts, devices for social control and management, that used the surface to organize data about locations, developing elaborate coding systems for integrating textual information and depicting physical features in detail. They took longer to refine projection techniques and uniform scale, using workarounds like writing the distances, once scale was not very much reliable. They had different priorities from Greek geographers and only stabilized uniform scale around the eighth century. To the Greeks, the emphasis seemed to be in exploration and in the strong combination with theoretical geometry that drove them to develop larger abstract systems.

Despite these differences, we can identify in both technological strands a tension between two uses of the grid: first, for mapping through landmarks and lines of distance between them, so the use of the grid would come as an aid for proportion, and the map would be a large collection of individual and relative measurements; second for mapping through geometrical projections, where the grids would serve as systems for point location by coordinates. From one process, position would be attained through the organization of many linear measurements, from the other, linear measurements would be achieved by projecting positions. In a sense, the history of maps can be seen as a history of transforming measurements into machines for more measurements, for more reproduction of

measurements, or it can be seen as the history of developing systems for relating textual information with a spatial display. Both are present in visualizations nowadays.

Therefore, maps are produced from a set of many measurements and/or projections of distances and/or positions, and are afterwards used to produce other figures, derived of intersections and sums, extracted with the compass and calculated with the scale. The measurements made initially are reproduced to be again verified and reproduced in new calculations. The map organizes measures but is also a device for measuring, it projects measures back into measured space. Here we have broadly the same idea we have discussed in the previous chapter: that graphic structures constitutes data in order to be reproduced and verified. The map is past course but also a device for future courses.

According to Foucault (2003), nowadays we have a substitution of extension by place: the place is configured once one specifies a point from where to act in the space. This new perspective has connections with contemporary fields such as demography, geography or logistics, that address the problem of the arrangement of things and how they are displaced. If in the Middle Ages, the place of things was for them origin and destiny, nowadays the many possible spatial arrangements set the stage for an endless and problematic rearrangement. Therefore, the issue of place is not solved in arranging things in such a way as to have space for everything, but goes beyond, encompassing also a management of the relations between things, their storage, circulation and many other variables. The forms that better relate to this perspective are actually the diagrams and mappings that account for places of importance and the relations between them, and not exactly extensions, surfaces and frontiers.

While considering contemporary cartography, November *et alii* (2010) describe the separation between physical and human geography, where the first is more connected to the description and measurement of physical features of the land, while the second focuses on organizing informations that include statistical data about the populations like revenue distribution or patterns of consumption. This second approach produces maps where the contours of land and seas, rivers and slopes, political frontiers and region perimeters would be used more like a diagram for the distribution of the main information. This latter perspective, already rehearsed since ancient chinese cartography, draws a connection with visual data analysis, like in Bertin's work with statistical maps. Nevertheless, as digital navigation gets more sophisticated, this diagram in which figures are displayed acquires much more plasticity in its

use, with zooms, distortions, pans and filtering, emphasizing a topological sense, like in network maps, that have no previously set reference system.

This leads us to the differentiation between the mimetic use and the navigational use of maps that is proposed by November *et alii* (2010): first, they argue that maps have always been an interface for calculations. As we already discussed, maps aggregate measurements and projections for new calculations. According to the authors, digital technology transforms cartography into a navigational platform, that aggregates databases, an interface for its treatment and recovery, a panel for the interaction with other users and a variety of possible outputs, one of them being print media. Moreover, all the paraphernalia of many actors, institutions, productive and publishing resources that is always present, even in the production of regular print maps, becomes more evident, more visible, has to be dealt with, in the realm of digital navigation.

Second, they point out that digital maps break more clearly the mimetic relation in navigation, because it becomes clearer that the user does not seek exactly a relation of similarity between the map and the territory to which it refers: the user relates to special points and figures that are relevant to his/her objectives. It is still a relationship of correspondence between representation and referent, but this time it is guided by the recognition of useful points that connect the two surfaces.

For the authors, cartographic traditions still inherit from nineteenth century paintings the need to put the map under the gauge of the values of mimetic representation. Mimetic representation works towards organizing the relation between two poles, the prototype and the copy, where this last one will always end up projecting a virtual image of what it is supposed to be representing. That way, the diagram composed in maps will project its lines back to the space they refer to and organize our experience with it in a regular, unified way.

For Deleuze & Guattari (1995, p.17), the notion of unity shows up only when in a multiplicity the significant takes over. In this case, the many experiences with extension, place and displacement are taken over and unified in a continuous space by regular projections. The unity always operates in the core of an empty dimension, supplementary to the one of the system that is at stake (overcoding). This takeover can assume the most varied forms, be it in the constitution of a language or code, of a territory, be it in more provisional processes, like an interactive path that organizes an overcoding by a specific access point. Maps, as visualizations in general, are in a sense an overcoding that stabilizes a complexity.

On the other hand, scientific images function differently: they should be evaluated inside a series, because its power as an evidence is not inside the image, in its constitution, but in the constant that is perceived in the passage between images. So there are much more correspondences being mobilized by the scientific images, but they come to light along series of inscriptions that are assessed over and over and progressively refined to highlight constants that must be measured and calculated. An image like this, when subtracted from its series, will go back to project, like images in general, a virtual image of the thing which it is supposed to depict. In this sense, it can even have aesthetic and rhetoric value, but will be disconnected from the value of evidence. For the authors, that is just what happens in traditional cartography when there is the impression that the map corresponds to a territory, that it is a decal of the territory.

So, as soon as the image is removed from the series, it loses its value as evidence and becomes mimetic, projecting a virtual image of the thing it depicts. Although this projected image looks like it is external and previous to the representation, it is no more than a projection of what is organized in the image. The notion of territory is always a virtual image, a projection of the mimetic perspective of the map. So we can say both territory and map are representations of a useful perspective, organized in accordance to the dynamics of the sensorimotor apparatus that, as we described in the first chapter, reinforces schemas that build a continuity between vision and pragmatic action. The map is extended into spatial displacement by simplifying sensorimotor schemas.

On the other hand, we understand that the experience of navigation, even with a traditional map, also involves, to a certain degree, a non-mimetic approach and a reconstruction, a structural game that reproduces – produces again – and reenacts the initial measurements. With digital technologies this is suitably equipped with specific features: once it is possible to navigate by layers of complexity and detail, operate panoramic movements and even distortions in its structure, digital navigation is recomposing a series. There is an exit from the mimetic perspective and, in this destabilization of resemblance, there is a precision gain and certainty in the measurements and positioning.

Also, this discussion, even if it is broad, addresses a specific conception of maps, the ones that refer to a spacial organization that is exterior to them. As we pointed out, many times we use the term map to indicate structures that aim at organizing just networks of elements, like many kinds of diagrams. For example, a map of allocation of resources of some institution

along the working week. There is not exactly a geographical territory being described in scale, but a distribution of concentrations in a grid of allocations. In this sense, the confusion we indicated in the beginning of this subchapter is significant: the term *chart*, which has the same latin root as, for example, the term *carta*, the Italian equivalent to maps, is both used for nautical maps and for diagrams displaying the most varied kinds of data. So we can see that a map is not only a device for displaying measurements of distances, but that, in a broader and more fundamental sense, it is a spatial allocation of lines, concentrations and directions, where distances multiply and may take on different measures, depending on the lines to which they are attached of related.

To Deleuze & Guattari (1995, p.16-17), a grid (like the ones built in maps) is a plan of consistency: like a field of forces, of lines of flight that always point to the outside, recombining until filling up the plan. The plan gives consistency and selects from the multiplicity, it is the space it fills and at the same time the projections to the outside. This description matches the main aspects we have been discussing about the grid in cartography: a game of lines. The grid, in this sense, may encase measurements, but will also provoke transformations in these measurements and will propose a consistency that is very different from the mimetic decal from a territory, denounced by November *et alii*.

What November *et alii* call a mimetic perspective, Deleuze & Guattari will call a decal. For these last, the map is something different, more productive: it would be open, always at issue, and may always be connected, relinked, restructured. The decal, on its turn, would transfer structures, would always be a reference to something, from where its stability would derive. A map is renovated according to the entrance point (and there are many), while the decal always returns to itself. A map is a matter of performance, while the decal always refers to some presumed *competence* (DELEUZE & GUATTARI, 1995, p.22).

So the map described by Deleuze & Guattari is very close to what November *et alii* elaborate on the subject of the digital navigation, while the decal is close to the mimetic perspective. From our point of view, visualization rehearses both perspectives in its performance, it may be at the same time map and decal. In this sense we may also say that there is a movement towards the abstraction of maps, not for rising them to an elementary or structural or even ideal form, but, again, as an interface for calculation and multiplication of variables in a plan. We believe that it is in this aspect that cartography lends its strength to the other visualizations.

So we notice that the possibilities for an exit from the mimetic perspective of maps do not issue from digital navigation: in fact, this innovation opens a window to the recovery of cartographic varieties and territories that do not fit in the standards of classic cartography, rooted in the mimetic reference to a external territory. Documents from the turn to the twentieth century (see, for example, WINKLER, 1901), describe the use of *stick charts* by native fishermen in the Marshall Islands. These were instruments for navigation, that at the time was made only in canoes. They were made from small sticks tied together, occasionally with shells, and, instead of seeking to represent the shape of the costs of the islands and their distances in the most precise way, their priority was aligning the variations and ondulations in the waters and their currents in the regions, and how they were reflected in the islands. The record and the knowledge of the patterns of waves was not primarily important for avoiding risks: the primary use of these charts was for orientation. The sailor would orient himself by these patterns of dispersion of the waves and the currents, recognizing them with the movement of the canoe. This technique was so refined that there was actually a classification of different kinds of these water movements. Another amazing aspect of these charts is that they were not carried in the journeys by canoe, but memorized and left in the shore. This leads us to believe that maybe, more than instruments of information storage and recovery that would project a virtual image, the stick charts were devices for organizing a knowledge: the lines in the sticks helping to compose an organized memory of the experience with a territory. What interests us most in these maps is that they are examples of a cartography that focuses in patterns that expand and aggregate, resurfacing a logic that is present in maps in general. In fact, some centuries of western mapmaking will go by before we manage such abstraction and economy of forms, and it will come only through many interactive layers, that alternate between depuration and complexity.

Despite being very abstract, the stick charts also differ from the kind of geometric abstraction developed in the west since the theoretical approach of the ancient Greeks towards cartography. These last developed a strong cartographic basis based on geometric demonstrations and projections that would bring regularity to a space that would always, by all means, be experimented cognitively in many different ways. We refer to Serres (2004) to advance that the price geometry pays to build an universal language is the fact that no form that is shown though it corresponds really to the thing it measures and demonstrates. Roque (2012a) also explains that geometry will seek to access the absolute that is encased in the forms, and that therefore, all forms are preexistent to their demonstration. So constructing a

form is in fact a way of understanding it. Back to Serres, we can say that geometry is always creating to itself a land to be measured, or many lands, given the variety of geometries that exist under a general impression of unity. From this point of view, every cartography does not cease to be the projection of a referent: one that does not correspond to it, but to a transitory presentation of a knowledge about it. Al last, the cartography continues to be a trace of the experience with a territory, be it physical or abstract.

This also points to another important issue we should consider, which is the relation between the dimensions of information and space. Again according to November *et alii* (2010), the classical idea of space was constructed with euclidian geometry, where, in a regular plan it is possible to position constant objects in different ways, without changing their constitution. Euclidian space is like a repository where objects can be moved without transformation and become calculable along all these different positions. This vision is similar to a world drawn on paper according to the principles of perspective drawing and, later, to projective geometry (MAYNARD, 2005). It is as if the world was duplicated by geometry and afterwards reduplicated back as a virtual image, a regular space that would stand for our world. In fact, there is not a sequence from one stage to another, or first one translation and then the other, because we constitute this projected image from the moment we see the world as space. Benedikt (1996) advances that information in space is space in information, taking into account that not only there is no ontologic anteriority of the territory in relation to the map, but neither of space itself in relation to information. Sensorially, space only exists as part of our experience when it is already and at once information, and vice-versa. We come to realize that a space that is external to any information is indeterminate and outside our experience, and that measuring and even cognitively processing space unfolds extension, in all directions. Each technique of representation produces, in short, a navigable space accumulated by all the maps we see as a mimetic representation or navigational device.

Foucault (2003) develops the concept of *heterotopia*, according to which there are other spaces that are not the space to which we got used to refer to. These spaces are like layers of experience that relate to the whole of the regular spaces. One of the characteristics of heterotopia is that it manages to superpose, in the same real space, several spaces and places that by themselves would be incompatible. The space of a theater stage and the cinema screen are examples in a list where one could easily add the computer screen or, generally speaking, all the digital devices, because it is through them that the informational spaces of our time infold.

This principle of superposition of heterotopias draws attention to the access in layers of variable complexity and scope that is characteristic of digital navigation as described by November *et alii*. It seems that, apart from accumulating many spaces in the same surface (like the cinema), visualizations accumulate spacial variables that can be mapped in many different ways, creating different spaces, different plans of consistency.

4.3. CONTEXT

Data is not useful in itself. It needs to be applied into a certain structure, either for feeding information systems, for communicating or for analysis. Communicating through data can be done in many ways, from direct description, to presentation in tables, to visual tools, while some analyses can also be carried out through abstract mathematical procedures. We would like to contend that the labor of turning data into something meaningful and useful is related to the search or the construction of a context *for* and *of* data, especially when visualization is used. As we mention specifically the contexts related to data and data visualization, we should also acknowledge that the concern about the hyperproduction of data through digital information systems often hides the fact that disaggregation in information and knowledge systems and the counter-effort of building context are not recent nor bred only in digital technologies.

In this chapter we will explore different aspects of the idea of context as it is organized and produced in technologies of intelligence (Lévy, 1997), from three main points of view: the context of origin, the context of use and circulation, and the internal context of relations that is created in any artifact of the kind. For this third part of our antiquing, that has already gone through discussing the experiences of time, space and now turns to the experience of context, we will trace some parallels between examples of cosmological representations and examples of network maps. The first ones will be related to the effort of representing totality in systems of equivalence and the second ones will help us to approach the building of context through partial sets of relation. Between the two groups there is a transformation of the position of the observer and of the procedures for producing or establishing knowledge. So we would like to contend that with the former we would have systems of unified reference from a totalizing point of view, and with the latter, sets of contextualizing relations with many possible entrances. We believe there is a productive tension between on the one hand the visual representations that aim at putting together a totalizing vision from the outside and on the

other hand the ones that are outlined mostly by intrinsic relations forming partial contexts. These are actually two complementary movements that happen in many levels, and their interaction is quite relevant and fundamental for our understanding of contemporary visualization.

Roque (2012) describes the space in aristotelian physics as a universe divided into two regions, the sublunar, where there was erratic movement and degradation, and the superlunar, where there was perfect movement, circular, eternal, constant. It is impressive how variations of this separation this still ring true in some popular conceptions of scientific activity: erratic movement for living bodies and social interactions, regular movement for celestial bodies and the concerns of physics. The place of a body would be defined by its essential qualities, so movement would derive from the tendency of a body to return to its essential place: this matches the importance of the idea of order in classic greek cosmology. Even though cosmology has taken up many facets, this effort of finding aggregating perspectives remains, by descriptions that can represent the universe as an ordered cosmos, a whole where order holds, as complex and hard to access as it may be. Geometry, drawings, schemas and lately visualizations remain as tools for building a territory for the calculus of the universe.

According to Roque, this principle of regularity and order present in aristotelian physics and the belief that the earth was the center of the universe demanded that the movements of the cosmos would be represented as concentric and aligned circles, even if this did not correspond to direct observation. Planets and other celestial bodies do not seem, to the naked eye, to be moving continuously, always in the same direction, and at times seem to close up or withdraw from Earth. Even so, many physicists, astronomers and mathematicians after Aristotle tried to depict systems that would convey the universe according to these principles.

Foucault (2003) proposes that, in the Middle Ages, space was seen as a reflection on divine order, having a stable order, organized in facing poles, like heaven and hell, sacred and profane etc. It was a scholastic translation of the classic greek cosmology. This space was orderly arranged, where every living thing, every tiny thing had its place or tended towards it, and every sickness or malefice would issue from drifting away from this place. Knowledge was also derived from analogy, and this was projected in systems of visual representation of what was known. The cosmos was represented by these same principles and cosmological representations were one of the main visual paradigms of this world vision. The world was

finite and there was a tendency to symmetry and regularity in astronomical representations, that became flattened and stable (see figure 4.3.1).

Figure 4.3.1 The Ancient and Medieval cosmos as depicted in Peter Apian's Cosmographia (Antwerp, 1539). It mixes astronomic observation, astrology and christian traditions. The text reads: “The scheme of the aforementioned division of spheres. · The empyrean (fiery) heaven, dwelling of God and of all the selected · 10 Tenth heaven, first cause · 9 Ninth heaven, crystalline · 8 Eighth heaven of the firmament · 7 Heaven of Saturn · 6 Jupiter · 5 Mars · 4 Sun · 3 Venus · 2 Mercury · 1 Moon”. Source: http://commons.wikimedia.org/wiki/File:Ptolemaicsystem-small.png

Roque (2012) describes how the Scientific Revolution along the sixteenth and seventeenth centuries transformed scholastic knowledge by rediscovering their ancient roots in the classic greek traditions. One of the landmarks for such transformation was the work of Copernicus, which is set in the beginning of a trail of many different versions of cosmological schemas that are tried out in the attempt of better matching representation and observation (see figures 4.3.1, 4.3.2 and 4.3.3.). We identify an effort to better describe the placement and movement of celestial bodies that will progressively put observation ahead of the principles of divine order. To Foucault (2003), the great revolution brought on by Galileo is not exactly the idea that the earth is not the center of the universe – at that time this had already been advanced by some astronomers – but the idea of infinity that it entails. Understanding the cosmos as something infinite breaks stable positions and mirrored relations, so the stable places of medieval times are destabilized, and we are thrown in the open field (SLOTERDIJK, 2011). Foucault (2003) describes this as trading location for extension, in a world indefinitely open where the place of something tends to be a point in its movement and things themselves become stable only by temporarily slowing down this constant movement. He understands this transformation is clearer from the Renaissance on.

Figure 4.3.2. Theoricae Novae Planetarum, by Peuerbach, 1488. The cover page displays the orbit of the Sun around the Earth. According to Roque (2012), Peuerbach seeks to account for details in observation (hence the width of the orbit and the relative position of the sun in it) while still keeping the geocentric model. Source: http://www.cbi.umn.edu/hostedpublications/Tomash/Images%20web%20site/Image%20files/P%20Images/pages/Peurbach.Theoricae%20novae%20planetarum.1488.title.color.htm

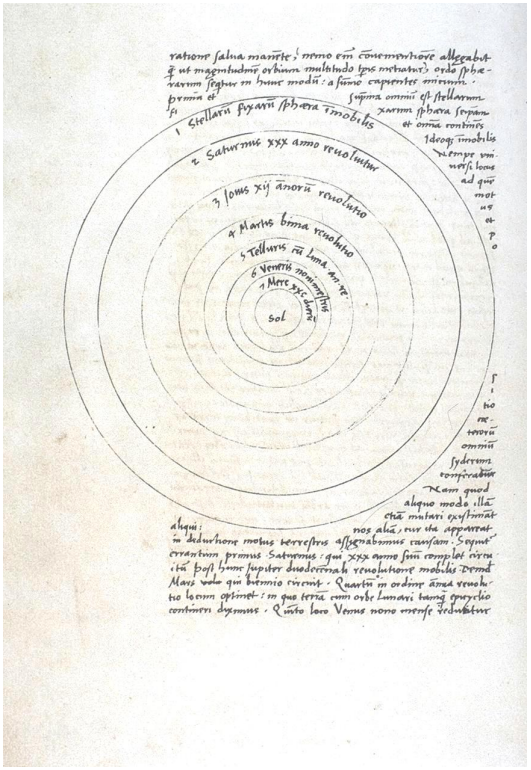


Figure 4.3.3. Heliocentric model of the solar system in Copernicus' manuscript. "*De revolutionibus*", p. 9 verso. The book was printed in 1543. Source: http://commons.wikimedia.org/wiki/File:De_Revolutionibus_manuscript_p9b.jpg

We believe a corresponding movement happens with the development of perspective drawing, from Renaissance on. For Panofsky & Wood (1991), perspective drawing turns images into windows to a projected space. The authors advance that the planification and regularization of pictorial space create a certain continuity with physical space. In a sense, the projected pictorial space, where arrangements of the physical space are depicted, projects its lines back into physical space when observed, because the bodily position of the observer suddenly gets transformed into a point of view upon what is beyond this window.

A bit along these same lines, but from a very different entry point, Latour (1985) stresses that the continuity between pictorial space and physical space in perspective drawing does not

entail, of course, an equivalence of both spaces, but allows for some exchangeability mediated by geometry and measure (see figure 4.3.4). Once pictorial space becomes uniform and discrete by the use of projective geometry, scale and proportions can be calculated. They are a regular set of relations that is linked to the relative distance between objects. So the geometry of pictorial space is aligned with the geometry of physical space, allowing for a network of correspondences. This creates a necessary connection between proportion and position that is a requisite for our analytical appreciation. At the same time, objects depicted become movable, they can be rearranged in the regular geometric surface without losing their structure.

Figure 4.3.4: Albrecht Dürer, Artist and nude, circa 1525. The devices and labor of building regular vision in perspective drawing. Source: http://commons.wikimedia.org/wiki/File:DURER2.png

Foucault analyzes the painting *Las Niñas*, painted by Velásquez in the first half of the seventeenth century (see figure 4.3.5): it depicts the painter himself, in front of a large canvas of which we only see a small portion of the back, given it is mostly out of the image's frame. Beside the painter, a domestic scene with the *infanta* Margarita and her maids is set against the sombre and tall walls of a large hall. It takes some observation to realize that Velásquez built this composition in such a way as to put the observer of the painting in the position of the model of the other painting he is working on in the image. It took us the hints provided by Foucault to understand that on the far and dark wall behind all these elements there is a mirror turned in the direction of the observer, and this mirror shows the image of the king and queen, posing for Velásquez. So this painting performs a series of movements: when most paintings would be depicted from the point of view of the painter, which would later be assumed by the observer, in this case the painter puts himself in the picture, operating a game of mirrors, in order to put the models in the place to be assumed by the observer. So, when some of the

people who are depicted are looking in front of them, they are looking at the models, not at the painter, not posing to be pictured. All this builds a complex arrangement of lines of sight that plays with the stability of the relation between vision and pictorial space that was organized from perspective drawing on.

Figure 4.3.5. Las Meninas or La familia de Felipe IV, pintura de Diego Velázquez (1656). Source: http://commons.wikimedia.org/wiki/File:Las_Meninas_(1656),_by_Velazquez.jpg *See Annex I, figure A1.4, for color plate.*

Moving ahead, Crary (2007), while discussing the transformations during the eighteenth century towards modern sight, will translate this in terms of a disaggregation between vision and the other senses, towards its specialization and abstraction. For him, this is part of an epoch where images start being mechanically produced that have no anchoring on bodily experience, like microscopic photographs, for example. Again, an analytical perspective is reinforced, as well as vision as if from a window, but here it becomes clearer that the image is turned into some sort of machine for seeing.

In classic cosmology there was no place for the point of view of the observer, the representations were distilled from sets of concepts, it was a view from nowhere. In

perspective drawing and in later astronomical schemas, the observer is entangled in the projection: even if the point of view is placed in a location that is impossible for bodily experience, it is a place relative to the projection. The central condition to all that is that our relation to space become optical, and that our vision become abstracted and analytical, so we can determine position and depth by evaluating and comparing relative proportions, angles and planes. Context is built by means of such positionings and conversions. In both cases, of astronomical representations and of perspective drawing, we see the visible traces of a passage between medieval analogic thought and the modern categorial, analytical thought. Nevertheless, as we discussed in the previous subchapter, geometry and measurements built a regular space to be measured and converted, but, as we aim to advance in what follows, twentieth century physics disentangled cosmology from totalizing visual representations.

Contemporary cosmology in physics relies on mathematical models that are often translated into visualizations, but have their foundations in measurements and abstract demonstrations. According to Novello (2006), cosmology was only accepted as a branch of physics in traditional scientific circles in 1964. The idea of studying an object that can never be observed (or represented) in its entirety, that can never be seen from the outside, since it is infinite, was faced with much resistance. It was not enough to establish, in the 1920s, that the universe was expanding: this scientific fact, by itself, would be applicable to the universe as a whole and would exemplify the possibility of of approaching it as an object of study, even in the impossibility of representing it in its totality. It was necessary to establish the measurement, in 1967, of the radiation of 2.7 degrees kelvin, that fills the whole universe, for scientists to ponder that, even in the impossibility of probing and measuring every centimeter of the cosmos, it could still be somehow studied in its totality.

But the plot thickens: physicists understand that there are four fundamental forces that govern the processes and the dynamics of every phenomena in the universe, but the only one that has a broader reach (supplanting atomic scale) and effect (considering that it affects even the same particles that serve its propagation) is gravitational force. Therefore it should be the main concern of cosmology. To develop a gravitational cosmology, it would be necessary to associate it to a unified space-time structure, that could be described by geometry. But, as gravity affects itself, it creates some sort of curvature of space-time, that distorts the regularity of the space-time continuum, demanding for different geometries to be applied in different fields. According to Novello, Einstein developed some theoretical reflexions on how to build this new geometry that would be able to describe universe as a whole, but it was left to other

physicists to prove that gravitational force indeed has the power to modify the geometry of space-time, and they have done it by calculation. It is interesting to notice that geometry tends to build diagrams in a potentially infinite space: that extends its lines indefinitely and in this sense tends expand beyond the representational space. Nevertheless, the extreme of representing all there is, that is, the whole universe, breaks its regularity.

When we consider the evolution of natural sciences and their many probes, from spacecrafts to microscopes, we realize there are infinities in every direction: in micro and macro scales, in the past and in the future, and also in the production of information about all these infinities, and the many appropriations and interpretations of circulating information. It is important to stress that, when we talk about context, we are also talking about the modulation of the limits of the dimensions of time and space.

An interesting metaphor that plays with both continuities, of time and space and connects it with the problem of information would be the fictional library of Babel described by Borges (2001), where hexagonal book rooms are connected with one another by mirrored halls, in which there are stairs that spiral indefinitely upwards and downwards, leading to other levels with the same structure. The library is so huge, that people are born, live and die in it without ever getting to reach its end, and therefore have it as their universe. They are referred to as librarians. All the shelves are the same size and all the books have the same number of pages and standards for text display, and use the same alphabet. Yet, the library contains all the possible texts, their possible versions and text commentaries, in all possible languages and some impossible, all the possible combinations of letters, and, to an extreme, all the possible gibberish combinations. So the situation is that, according to the narrator, the books that carry actual readable text or text bits, regardless of their language, are hidden between a huge amount gibberish letter combinations, to an extent where a person will probably go through his/her whole life in the library without reaching one single book that offers meaningful content.

This library turns potential combinations of discrete elements along regular structures into architecture and bibliography, but at the same time, in its extreme, it dismantles literacy. We see it is a metaphor for intertextuality… without context: even if we can imagine secret and codified connections between the books, and also imagine, with the narrator, the existence of a full and accurate catalog of all the library somewhere along these many shelves, which would equal the word of God, such connections or indexes are not accessible to human

experience in this dismaying universe. To bring this architecture into the sphere of human existence it is necessary to build layers of interpretation that would generate more text and images, but also would allow people to navigate, debate and exchange, it would build context and open space for richer human subjective experience with the books in the library. The books about books have their own intertextuality but produce intertextuality between the books already given in the library. We are talking about creating knowledge systems by building texts about texts, bibliographical interfaces, and, to an extreme, metadata.

The main character and narrator of Borges' The library of Babel explains different theories about whether the library is finite or infinite, and at times these theories resort to statistical arguments, like the idea that there cannot be an infinite combination of letters, considering that the books have all the same shape and length and use a limited set of base characters. At times, they resort to spatial or cyclical metaphors, like the final one, upon which he settles: the Library is unlimited, but periodic. If an eternal traveler should journey in any direction, he would find after untold centuries that the same volumes were repeated in the same disorder, which would then reflect an order, the Order (p.36). From our point of view, this would be the cosmological solution: a tautological formulation that entertains the possibility of an eternity (as in the eternal traveller) to experience an infinity, and takes on a view from nowhere, aiming at totalizing representations.

There is also a connection to be made between the extent of information, that is potentially infinite, and the infinite space of the universe. Of course, this library can also be taken as a metaphor for our physical universe, where the things that are at our reach may not give us the answers we seek, no matter how much we keep extending our perimeters and diving into detail. So another solution would be the modern one, of exploring, probing, building progressive layers of classification between the books, and mapping halls and writing compendiums that, despite all efforts, would always be incomplete, but at least would be at reach. It is a critical, analytical perspective that allows for accumulation of knowledge. Each would gather a full collection and, in the same way as with the examples of perspective drawing and modern astronomical representations, each would contextualize the reader in the universe it creates, guide him through a central, top-down analytical point of view.

There is yet another path, that we relate to a contemporary perspective, which would be to find connections or possible references between the books, therefore roaming the library oriented by such references, experiencing point by point a non experienced intertextuality.

This point by point route is exemplified by the narrator of Borges' tale as a fair but impracticable alternative for understanding the books in the library. While the books in the library themselves, consisting of paper binds and not having any apparent order despite the countless number of shelves, could not be mapped into this experienced intertextuality, or only in a very restricted sense, if they are translated into data or inserted into information systems, such connections can be built in a broader sense.

In the visualization in figure 4.3.6, we see an interesting effort of converting a large scientific compendium into a visual landscape: according to Börner & Polley (2012), it displays the text of four volumes of the Henry Smith Williams's *A History of Science* (1904) in the ellipse that encompasses the word cloud. The full text of each book fills a quadrant of the ellipse. The preface of each book is displayed in the correspondent corner, while the introduction to each major school or movement is displayed in columns that radiate from the center, also matching each book's quadrant. The word cloud in the center represents the terms found in all the four books, where the ones in larger font size would have occurred more often. The words towards the middle would have occurred throughout the text, while the others may float nearer the bits of text where they were most common. The words in red are the ones with capitalized first letters, typically proper nouns like names of places or people, and will in general float to the borders.

The first thing that strikes us in this visualization is that it contains the whole content it aims to map. In this sense it looks like a map the size of the territory, with the disadvantage that the text is harder to read in its fullness. But it also makes us realize how much linear text is non-visual and, well, hard to grasp in its objects and relations. In the case of the library of Babel, there is the added problem of the broken literacy, so bringing the content of those books into human experience is difficult. Nowadays in western societies we have the issue of excess and disaggregation, that causes a similar problem of making it all, in practice, a bit unreadable. But, as a basis for this visualization, we have a classic compendium about the history of science that is supposed to be, by itself, a roadmap. It is a reference work already packed with indexes and stabilized categories, typical of the modern perspective over knowledge. Nevertheless, this visualization offers a complementary path to the content of the book by correlating its classic structure (the books and chapters, for example) with what was found with word co-occurrence analysis. The structure of the book becomes the starting point for what the authors call the visualization's reference system, while the word cloud in the center will allow for a quali-quantitative evaluation of the compendium's content, taking into

account variations of density on the subjects and so on. So this visualization is not like a map that is the size of the territory, but a map that turns linear text into a visual territory through indexes derived from its intrinsic components. The same way as the continuity of printed linear text is mapped by discrete elements, we should also remember that Foucault (2003) posits that nowadays our idea of space does not emphasize extension: space is seen as a collection of many places. And a place, in this sense, takes form in the relation of many locations, it comes from assuming a point from where one would act in space. Space and also text or information get to be composed as a set of variables, subject to endless managing and rearranging.

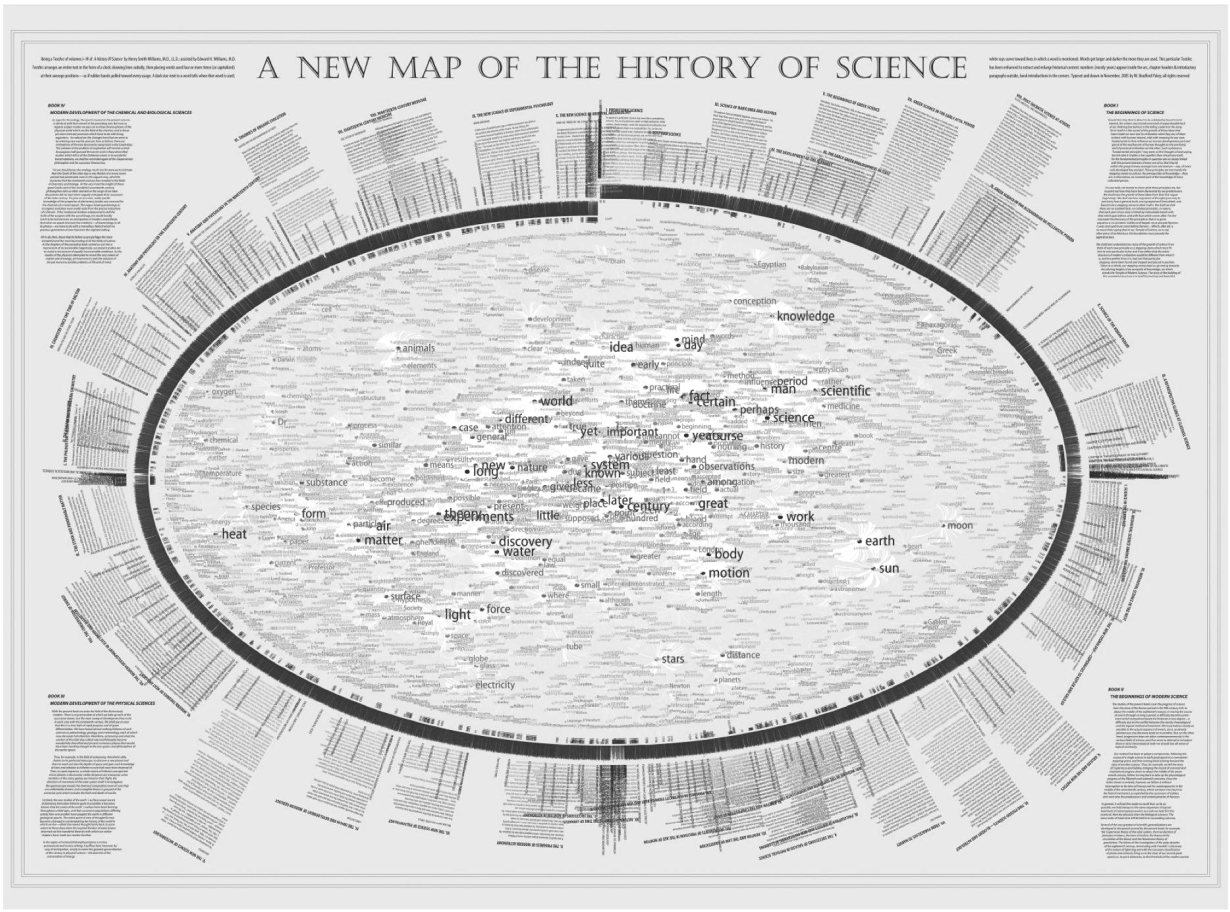


Figure 4.3.6: *TextArc visualization of The History of Science* (2006), by Bradford Paley (apud Börner & Polley). See Annex I, figure A1.5, for color plate.

In an age where we have unparalleled access to information, the quest for stable and totalizing representations recedes into the background, because, like in the library of Babel, the whole of the informational sphere exceeds our cognitive capacities but, unlike Borges' tale, it is always shifting. Börner & Polley (2014), for example, refer to a study where it was found that, as the

internet grew and more cutting-edge scientific papers became available online, a group of scholars from an american university started citing more from their own community than from all the many available sources online. So, just when we expected that the internet would cross geographic barriers, scholars tended to reinforce local ties. Likewise, context, rather than overarching order, becomes central nowadays.

We understand that network maps are among the forms that best represent this attitude, because they map out landmarks and their relations, not exactly privileging extensions, but relative positioning and density. A classic orientation about networks in information design is offered by Bertin (2011), who classifies kinds of graphs precisely by the relationships between base data: for him, the graph is a network when correspondences can be established between all the values in the same variable (or, in the terms of the author, all the elements in the same component). For example, the relations between elements in an organogram or a treemap. As Börner & Polley (2014) explain, they do not have a preset reference system, like many other graphs: it will be defined by the choices made according to the data available. So these visualizations are very close to data, in the sense that they do not derive of much pre-interpretation, and leave much space for interpretation and further refining in other visualizations or views that might be produced from it.

Network maps are heavily used in digital methods of social and literary research, in the set of fields broadly named the digital humanities. It is related to statistical approaches like in bibliometrics and scientometrics, because it emphasizes data-points and their connections, in methods like word frequency and co-occurrence analysis, and to the composition of an intertextual context by granular occurrences in the text itself. In figure 4.3.7 and 4.3.8, we see a network map built from Wikipedia content. It considered, in all the entries from the history of philosophy section, the relations of influence between philosophers that are shown in the right-hand box (see figure 4.3.9). So each node in the graphic represents a philosopher, and the lines are relations of influence. The bigger the node, the more influential the philosopher is. Nodes get to be closer from one another if they have similar relations of influence.

The title of the graph might be surprising: “Graphing the History of Philosophy”, because it might seem strange to imagine a history as a network. Indeed, the graph presents several different philosophers, collected along the whole history of western philosophical thought. Also, if the relationships of influence are to be considered something temporal, that would have at least partially a chained occurrence, we can think of this graph as representing a

history. Nevertheless, it is a very specific approach on the idea of history, because influences are not exactly seen as the movement of thought, concepts and theories transforming themselves from one into the other, as a result of continuous dialog and exchange. Influences are seen straightforward as connections, and not transformations, while philosophers are subsumed into quantitative and relational entities, like indices, more or less influential if they influenced more colleagues. Nevertheless, by losing this deeper dimension of the transformation in time, one can have a pretty clear and revealing vision of the context of knowledge these influences develop.

With all the details presented in the map, it is possible to evaluate the categories and attributions made by the community of Wikipedia as a connected whole. This allows many different entry points into the main theme and builds a new layer of appreciation. For example, by coloring the different philosophical traditions, one can discuss the exchanges of influence between them, and even elaborate on the adequacy of such philosophical divisions. On the other hand, it is also possible to reassess the links of influence of Wikipedia themselves, that were built mostly from a top-down perspective, that is, based in critical, categorial appreciation, and now can be evaluated from a different point of view. For example, Adam Hogan, in his blog Design and Analytics, brings up an interesting discussion about this network: judging for the centrality and the size of the node of Hegel, he would probably be the most influential philosopher in history, which seems curious, considering the fundamental place ancient greek philosophers like Plato have in western philosophy. So, is he really king? Moreover, did the editors of Wikipedia themselves had any previous consideration of how this aggregated result would look like when they went, point by point, defining the influences of each philosopher? Probably not. It was only visible… well, by visualization, by assembling a visual context to these scattered bit of information.

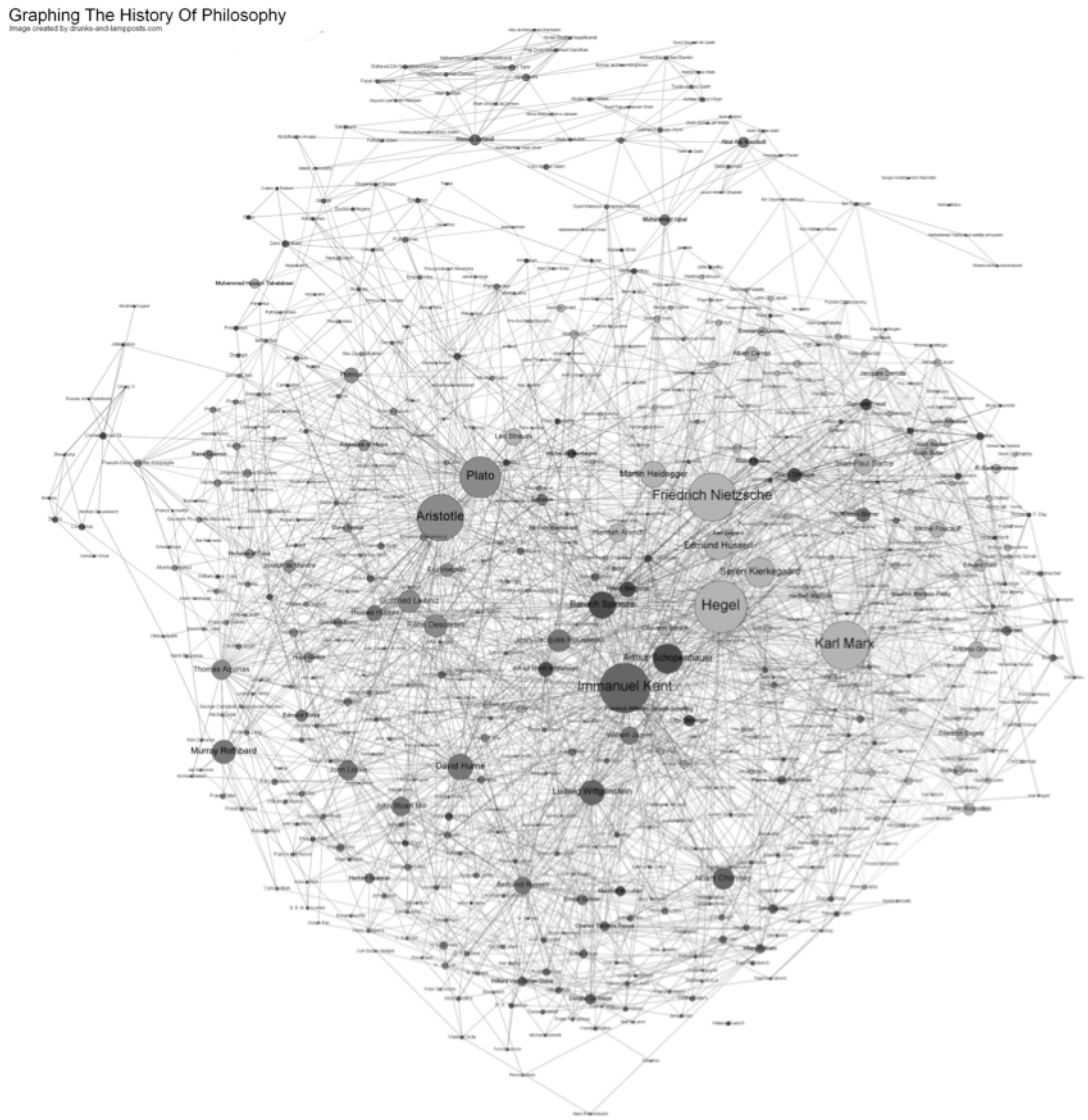


Figure 4.3.7. *Graphing the History of Philosophy*, by Drunks and Lampshots. Source: http://www.coppelia.io/2012/06/graphing-the-history-of-philosophy/
See Annex I, figure A1.6, for color plate.

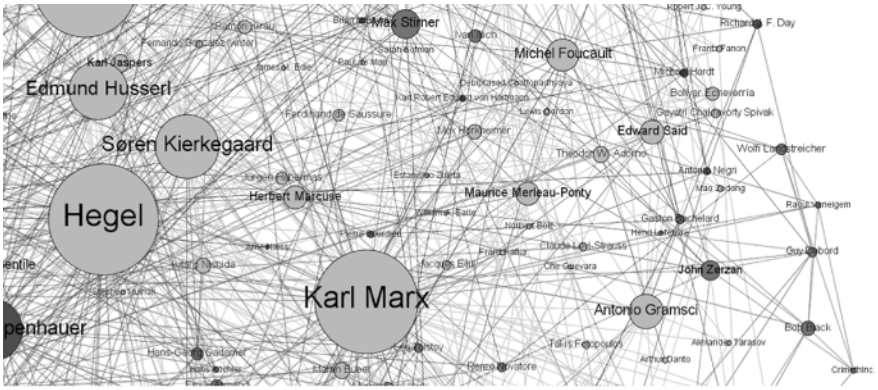


Figure 4.3.8. *Idem*, detail.

Of course metrics like that, that analyse content by the bulk and do not necessarily involve a previous critical work and stabilized categories, will in general generate some resistance. Some fear they might build a reductionist approach over knowledge, literature and academic production, by its quantitative aspect. Nevertheless, we believe both perspectives may be used to check on one-another, like in the case of the network on the history of philosophy according to Wikipedia.

Figure 4.3.9. Detail of the right-hand box of the Wikipedia entry on Immanuel Kant, showing the influences of the philosopher and the ones influenced by him. Source: http://en.wikipedia.org/wiki/Immanuel_Kant

All of these three historical variants carry a certain concern for totalization in representation, even if at different emphases and with varying translations of the idea of totality. They will be part of different cosmological views, even if, for example, thinking in classic cosmological terms may not be seen as something current in contemporary western societies. This is because building context entails a modulation of the limits of the dimensions of time and space, of framing and scale from the most particular and local context of experience to widening contexts towards unreachable universals and totalizations. And because we want to highlight this modulation that is key for the experience with contemporary visualizations, we will avoid opposing contextualization and totalization.

In a more fundamental approach, the building of context is deeply entangled with issues central for communication and language itself. Bertin (2011), for example, talks about defining a code, a system of signs that can be coherent along the graphics, so they can communicate clearly and efficiently. Every context functions as a plane of reference for cognition and for communication. Each visualization uses and develops visual codes and languages that are crucial for its construction and reception, and can build the character of the relations it displays. While some authors will advocate building these sets of visual signs from a perceptual basis, in order to avoid cultural biases in their interpretation, we would like to propose we look a bit more attentively to the processes of forming those contexts, that entail material from many contexts.

So once we want to discuss the building of context through devices of information, we have to acknowledge that the matter is manifold. In the specific case of visualization, for example, we can discuss the building of context while focusing on the many transformations of data and its aggregation towards a visual landscape; or we can privilege the associated communicational function of visualization, emphasizing its linguistic or semiotic dimension and the development of a system of signs that co-creates the insights derived from it. From this point of view, the context created by a visual display of information is understood as a field of enunciation. While both dimensions co-occur, we can identify the emphasis in the first one with data analysis and in the second one with presentation. Of course, there is no analysis without a concern for the communicational aspects of the visualization, as much as there is no presentation without visible relations between data.

In both dimensions, we propose three main aspects for the idea of context, one pointing for the past, one for the future and one for the insides of visual displays. Again, speaking from the point of view of data visualization, first we would have the context where the data was produced, that will give us hints about its coverage, possible biases etc. This part would also include the standard graphic structures and features that, once used, will guide the fundamental relations and proportions from which its visual language will emerge. It points to the past, as in the context of origin. Second, we would also have the context of use that is intended for the objects that use or are based on this data, be they software or graphs, which will entail critical concerns to HCI and information design, for example. This points to the future. Last, we have the context made visible by displaying data points and their relations and proportions, the display as a context in itself. Of course, as we have seen in the previous chapters, these three aspects will be constantly at play in the work with visualization. In fact,

building visualizations entails necessarily a concern with them, for these contexts will eventually allow us to refine the difference between information and noise at each situation.

We can widen these three aspects if we are to discuss the importance of building context in other technologies of intelligence (LÉVY, 1997): First, context as place of origin is related to where and when a certain text was written, or picture taken, or, for that matter, how a certain dataset was extracted or produced. Thinking about the context of origin of things, in this sense, involves relinking them with the socio-technical networks from which they were bred, towards developing a wider understanding of the tensions and agencies involved in their production. These networks include, of course, other objects, people and institutions. It also tends to emphasize a certain stability of these devices, that are taken as well defined objects, that are results and agents of this wider context, to which connections are traced mainly according to critical interpretations of their content. As with data visualization, this aspect of the idea of context turns towards the past. Second, there is the idea of context as related to where and when such devices may circulate, where and in which situations they could be used or be relevant. These appropriations also relink them to wider socio-technical networks, that include books, people, collections, institutions, technical infrastructures, so on and so forth. It points to the future. Both of these aspects are talking about fairly well defined objects whose story and uses can be traced and evaluated by a critical and interpretative approach towards their content. They are talking about intertextuality, but being mostly from a critical point of view, it will be highly penetrated by relations that are exterior to the content, building a context *for* it.

Speaking from the point of view of the linguistic dimension of the idea of context, we can also pull the three aspects we outlined: first, we have context as past, as the heritage of the visual structures that have been stabilized across much usage, from where the general features of the visual display arise, together with the assets and characteristics of our visual apparatus, that will organize the whole that is perceived in a certain way. Second, we have the related context of the circulation of these signs and structures, how they are interpreted and appropriated. And last, we have the display itself as a sort of context, that is the one that is more valued by Bertin, where the interaction between signs and graphic variations unfolds new possibilities for interpretation.

These contexts are always present: for example, if someone decides to apply scientometric methods to a specific collection of a library, say, the collection on astronomy, there will

always be some tension between the bottom-up categories the researcher wants to produce, and the top-down, critical categories of the scientific community and of the institutional priorities of the library that gathered that specific collection in a certain way. So both aspects of context are talking at the same time about intertextuality and about institutions and social networks of belonging and appropriation. In figure 4.3.10, we see an example of a visualization that is closely related to this perspective: the history of science fiction strands are mapped out in an elaborated hand-drawn visualization that summarizes and contextualizes the result of an enormous amount of work in literary critique. It aims at being exhaustive, and the strands that are not represented were not considered relevant for the story being told.

On the other hand, the third aspect of the idea of context is related to emphasizing the constituting relations and building the context from there. It involves a bottom-up approach, like the references from book to book in the exploration of Borges' library. It involved also laying out a network of connections that does not necessarily match institutional borders, for example. The focus is on building a universe from its intrinsic relations, even though classification will occur, and this intrinsic aspect of context will always influence and be influenced by the others. In this sense, the first two aspects of context are somehow semantical, while the third is more grammatical.

In the first two aspects of context (of origin and of circulation), the limits of the representation would relate and be outlined by external references, be them collections, institutions or predefined categories and critical standpoints. Therefore, there is an effort to account for the whole of this space that needs charting. On the other hand, when the context is built through deploying relations, the limits of the representation are outlined in the exhaustion of those relations: the network grows until all relevant relations are traced.

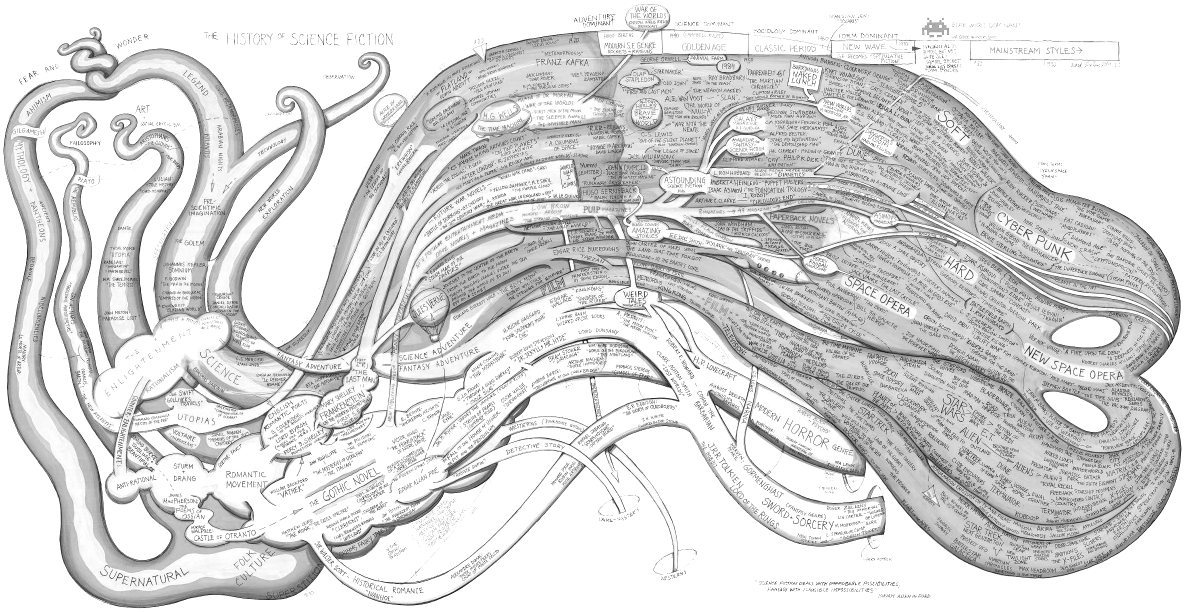


Figure 4.3.10: *The History of Science Fiction*, v.1, by Ward Shelley. Source: http://www.wardshelley.com/paintings/pages/HistoryofScienceFiction.html
See Annex I, figure A1.7, for color plate.

Nowadays, the movement that recedes and encompasses more, related to totalizing views, and its complementary movement, of closing in on the detail, become more and more a modulation of transitory and utilitarian perspectives in interfaces of computer software. More than aiming at representing the totality of life's complexity or establishing our incapacity for that, the necessity of creating relevant partial contexts or localizing entry points from where contexts may be organized becomes the central concern. Information is excessive and scattered, so it is not enough to filter the relevant parts: it is necessary to create relevance by contextualization. Each map, diagram, page or navigation menu intends to be a general and useful schema for some set of information and opens the path for a game of approximations, distanciations and localizations from which contexts may be organized and experienced.

4.4. LISTS, TABLES AND GRIDS

This subchapter closes the path we have been threading with our antiquing in the previous subchapters, by highlighting the presence and role of these three objects: lists, tables and grids. They are constitutive and effective across many technologies for production, storage and circulation of data and information, and therefore connect visualizations to wider traditions. Since the beginning of this large chapter we have been emphasizing the role of

such structures for creating discrete and interchangeable objects, but in the pages that follow we would also like to advance that an important feature of these is their interchangeability. The fact that more can be derived out of discrete elements once they are arranged along different structures is crucial for understanding the role of visualization nowadays.

4.4.1. Researching with lists, tables and grids

According to Bottéro (1995), lists are in the roots of phonetic writing systems. He takes special interest in the emergence of written language in Mesopotamia, because in this case the emergence of phonetic writing from ideogrammatic writing occurs inside the same society's history, making some developments more evident. To him, ideogrammatic writing is based initially in the evocation and the recognition of concrete objects of daily life, and always returns to an experience or knowledge already given, so it was inappropriate to elaborate something new. Ideogrammatic systems pose a higher demand that all the participants be immerse in the same concrete and coincident context, together with language itself. Therefore, to evolve it was necessary to adjust written language to spoken language; to adjust written record to words and not to things anymore.

Even before phonetic writing, the use of ideograms goes through some changes that will lead to phonetic writing: each object begins to represent not only an object in the immediate reality, but the word itself, its sound, so purely phonetic combinations begin to emerge, and the drawn figures become more disembodied and universal. It is therefore in this sense that written language is simplified and becomes discrete: in the sense that it is composed by combinable segments that are also independent of objectified references. Bottéro (1995) calls our attention to the social context of the time, explaining that there was then, in the region in question, the coexistence of different peoples with varied tongues, that were in intense commercial exchange. This must have augmented the pressure for the development of a writing system that was more abstract and functional, as well as an emphasis on data records, for managing sales and stocks. We have in phonetic writing an instrument that is more suited to the accumulation of messages, data and knowledge, that will improve a more broad and detailed approach, because it is more context-independent and exchangeable.

We understand that these developments also as an approximation to the contemporary idea of data, especially as it came to be applied in the fields of computation and in the management inside information sciences: one starts to produce knowledge from discrete elements which

demand the development of structures for recording that are capable of building more aggregate context of information. This goes for syntactical structures and grammar, and for graphic structures as well (GOODY, 1997).

On the other hand, this whole discussion is also related with the roots of scientific and literary traditions, that demand, each in its terms, the accumulation, maturation and interconnection for progress. As Latour (1985) puts it, one of the biggest challenges or goals for techniques of inscription in general is the production of what he calls mobiles immuables: knowledge objects that are stable, discrete and can be transported between contexts without losing utility or applicability. Bottéro (1995) speaks of lists as science: if language now is composed of discrete elements, the structures in which these are recorded grow in relevance in their application. He brings up mesopotamic boards where terms were organized spatially each in relation to the others, in a sort of classification and ordination: lists that organize a set of information. What this example shows is a radical change of approach regarding written record. If before the focus was mostly on making reference (with all the richness this might entail), with these lists in phonetic writing a new set objectives emerge: to know, to clearly register, define, situate, to classify and understand… these are concerns that become possible only through writing (BOTTÉRO, 1995, p.26).

These considerations make clear that the words, then, begin to refer to mental creations, to categories that gather different objects around the general accumulated knowledge about them. The lists described by Bottéro are not only lists of stocks and orders, like the ones whose previous existence has been documented. These specific lists called the anthropologist's attention because they are reference works, in the sense that they catalog and classify objects. Therefore, they work over language, return to language itself when aim at composing an aggregated and logical description of the world. This description is also utilitarian, because it aims at a practical use and dominance over worldly things. So we have the possibility of perceiving the set of things that are represented and sorted and also to understand its broad extension as categories, which is unfathomable by concrete means. The ability of speaking of a set of things without necessarily possessing them, not even in perceptual terms: lists are indeed powerful devices.

Sometimes it is hard to recognize when a list turns into a table. While Bottéro highlights Babylonian clay tablets that show an emergent science of the list, Gleick (2011) will point to yet other tablets from the same civilization and describes them as tables of instructions and

values that were later considered to be the first rehearsals of algorithms. Indeed, tables are similar to lists, but they have an extra dimension that brings on a qualitative change: they allow for metadata, that classifies records, and also allow comparisons and transformations in more directions.

While discussing the example of classical ethnographic methods as applied to the study of african communities, Goody (1997) cautions to the fact that lists can also work towards a reduction of the knowledge that is gathered with observation. That is a graphical reduction, where broader and more fluid concepts of spoken language are organized and narrowed in a structure that imposes order, turning them into discrete exchangeable objects. When the ethnographer gathers material through interviews and oral accounts, the organization of the record in a list entails a distancing of the more commonplace circumstances of speech. Categories are, at once, more visible and more abstract (p.150). Especially with tables, that can be taken as bidimensional combinations of lists, they smash accounts into discrete pieces to recollect them in a spatial structure that can be read in different directions, offer comparisons and reveal relations. There is a matter of continuity and discontinuity at each transformation.

Goody's concern is that the systematization of the records of interviews and direct observations might contribute to fit findings in ethnocentric perspectives, while disembodying and restructuring their original meaning. Nevertheless, he understands that the main goal of these procedures is to grant some sort of reproducibility to the data, which is a central element of all systematic development of knowledge. The term reproducibility here is not used in the sense of the ability of generating many copies, but in the sense of being able to produce again the same results. Goody uses the methodological notes and bibliographical references of academic papers as examples of this kind of reproduction: for checking sources, for showing how one can perform the same study again and reproduce the same results. Being able to reproduce is being able to verify (GOODY, 1997, p.118).

As the systematization of information becomes more complex, lists are superimposed and combined into tables that relate term by term. The frontier between a list and a table can sometimes be unclear, especially because they are interchangeable, meaning that the same set of data can be made into different lists or into different tables, depending on criteria adopted, which will, of course, change the possibilities for deriving information from them. It is only that tables highlight non-linear comparisons and relationships, and work bidimensionally, that

is, in two different axes, while lists will highlight ordering and work mostly in one dimension, one axis.

With lists and tables the relation to language and text is quite evident: we could easily follow up on Goody's discussion about the systematization of ethnographic accounts, and talk about these accounts as narratives that are mostly linear and get broken up into concepts and categories, to be then ordained in lists or organized in tables for correlations and comparisons. Pretty much like literary text can be put into databases according to keywords and metadata that are, in sum, also organized as tables. On the other hand, there is also the aspect of time: lists are, in a very fundamental way, a structure that is going to turn sequence into an issue: how should entries be ordered? Which should be the first ones to be read? What do we discover or highlight by trying different criteria for ordering? Also, time is an important aspect when we think about lists or tables as sequences of events: again, there is the distilling of discrete elements, now in the form of events. It is enough to think about monitoring systems, that will align certain points of data to certain points in time. Should we present the events in chronological order? Or reverse that order like we do in blogs? Does the list end or does it continue indefinitely?

Grids, as we intend to discuss them, are instruments for slicing and scanning surfaces. A wonderful example of the use of grids in research is given to us by Latour (1999) when he describes the work of a group of researchers that study variations in the soil of the fringes of a forest. In his account there are mainly two grids: the one that is traced through the forest, slicing its surface in order to guide the sampling of different parts of the soil, and the grid of the podofil, a standardized suitcase where these samples are organized in small cubes, according to their original position in the forest. We consider this a beautiful image of what digitalization does: it gathers points, samples of the mess and complexity of our forests and makes them comparable, exchangeable, transportable. The forest is, of course, a much complex and confusing environment, but those samples gather what is relevant, for the purposes of that specific inquiry. Putting them in another grid (as in the podofil) will keep its reference to the grid in the forest, but can also make room for other arrangements. In this sense, computer scanners are grids for capturing color points, as are computer screens, for organizing color points according to data and code. Grids organize maps, floorplans, diagrams, newspaper pages and site pages. The cartesian plane can be seen as a grid. Tables can sometimes be seen as grids. Grids are interchangeable with lists and tables, it will depend on the goals we have for the data.

Until now, we have been talking about data and graphic structure mostly from technological and scientific points of view. The mesopotamian boards, approached mainly as taxonomic works that, as said, return to language while working over categories, the ethnographic tables and lists, as analytical instruments for field records, the grids that organize the sampling of forest soil and further sample analysis: these are all instruments for data distilling and analysis, for the production of knowledge. As Latour (1999) puts it, all of them work towards progressive transformations for more compatibility and standardization and less locality and materiality. If we focus on graphic or visual structures as we have been, we understand that, at each step in the sequence of transformations Latour describes, there is a movement of fragmentation, followed by the complementary assemblage of a new continuity in which structure plays a meaningful role. It is meaningful in the sense that it opens the path for multiple combinations that trace relations and move towards a narrative structure.

We would only like to point out that, as we have seen in many examples along our antiquing work, the transformations towards more compatibility and standardization do not align lists, tables and grids progressively. In the very case of the podofil, the forest is translated into a grid for the samples, the samples organized in a correspondent grid, that which may turn into a table of values attributed for the samples, and that may, on its turn, be organized in another, different grid, that will organize a representation of a transversal cut in the land, for example. Lists, tables and grids are translated into one another all the time, and there is not a necessary, progressive order between the three.

4.4.2. Mediating with lists, tables and grids

Lists, tables and grids are also present in everyday life in a surprisingly ubiquitous way. For Latour (2012), we live in a world populated by socio-technical objects, by networks of agencies between persons and objects where everything is far too narrated to be only the object of physics or to be considered purely natural. Either in everyday tools or in narrative forms that explain and describe our daily experiences, there is a growing penetration of techno-scientific content and procedure for our interactions, through digital technologies. Those tools and narrative forms will also unfold as graphic structures like lists, tables and grids, because exploring these debates or using these tools will demand that they be represented.

This phenomena, of the penetration of techno-scientific procedures and content in social life, might be described in media studies inside the scope of technical mediation. It highlights the growing integration of technological procedures – more specifically, digital media – with cognitive processes and with the production of subjectivity through mediation. Once our language and interactions are deeply integrated with devices of technical mediation, more about these processes becomes traceable. This entails another aspect, pointed out by Venturini (2014): there is a growing formalization of different aspects of social life, which means that interactions that were not accounted for, that were out of chart, become traceable and potentially the object of analysis. This formalization occurs specially through digital tools, such as social networks, when objects like lists, tables and grids are not only in their surface (in the graphic interfaces), but are constituents of the technology itself. These objects are in the roots and branches of digital technologies.

So there is this coupling of techno-scientific procedure and content with digital tools that mediate our daily experiences, work and social life. Of course, this will also highlight the importance of techno-scientific controversies in general public debate and citizenship in general, as well as the political and social aspect of the making of science and technology. For the purpose of this work, it is important that we look closely at those processes of mediation, at the medias themselves. It should follow up on that same double movement of fragmentation and reassemblage that we outlined previously, but this time considering processes of mediation and media devices.

This double movement is as relevant for graphic or visual structures as it is for the supports themselves. That way, scrolls of parchment get broken down into pages and reassembled in codexes, which will give place to books with numbered pages, tables of content and several kinds of indexes. Before typesetting, cursive text itself by the ninth century will be broken into separate words and and sentences, through the introduction of spaces, the difference between uppercase and lowercase letters and punctuation, followed later by paragraphs and different levels of titling, marking their internal hierarchies and groupings. Typesetting will break up the continuity of the lowercase cursive text and allow for the emergence of the press. The sequential and one-column grid of the classic book will, in the nineteenth century, inspire the graphic organization of the first newspapers, to be later broken down into a grid of several columns by the turn to the twentieth century and, halfway into this same century, organized in modules that group specific content and span across multiple columns, over complex grids. With this, and several layers of titling and visual cues, the reader will be able to get a grasp of

the content's structure at a glance, and, by given graphic emphases will explore the page in his/her own prefered order. Hypertext will take these same modules, indexes and multiple references to new heights, bringing content closer to data tables, for it growing modularization and interchangeability. Data tables are, after all, one of the basis for computation in the databases, distributed with computer systems just a few decades before.

Of course, in many points the story we just told will not have such tight concatenation, that is, will not be as fluid. This is especially true for the first transformations in writing systems: of course there were, back then and also nowadays, several different systems at hand, and most of them would not display the continuity of elements of medieval cursive writing, some of them would not even be phonetic. Also, the passage of the scroll to the codex is only central when we consider european practices inside the church at the time. As we move towards the twentieth century, transformations become more standard and unified in western countries. This is probably one of the consequences of what Latour (2013) calls the modernization front: progress and standardization for efficiency are part of this logic that, of course, also comes with free commerce and heavier exchange between nations.

Nevertheless, it is important to highlight that at each transformation there is more content and accumulation, more need for orientation aids, more decisions that are left to the reader, and there is more non-linear exploration. Lists, tables and grids will also be basic objects that will organize these different structures. In fact, looking for and revealing their presence helps us realise how they graphically organize passages between narration and data, a phenomena we think will be more evident in the content of the medias.

Lists, tables and grids are also, of course, everyday objects. In The Laws of Simplicity, Maeda (2010) advocates that simplicity should be a principle from our daily lives to business strategy and product development. When talking about the second law, Organize, he highlights the importance of the tab key, even in traditional typewriters: a "generous sprinkling of tabs" (p.16) creates tables, which makes categories come to life. Categorization, in terms of inscriptions, might be summed up to adequate spacing: "In the medium of text, tabs break up the linear space of a document such that the paragraphs can stand out as the organizing principle."(p.16) And we do need visual organizing principles.

There are lists, tables and grids in our refrigerator doors – so we do not forget to buy the groceries we need – and in our table calendars. Our smartphone apps are filled with them. From our drafted lists, where we rehearse many possibilities to organize our daily life,

ordering and reordering priorities, we come to the lists and tables that are in the base of computing systems, through a huge network of practices and technologies that structure our knowledge, many of them older than digital tools. Likewise, as systems become more sophisticated, there is a growth in the variety of presentations and uses of those basic objects we have been talking about, in order to give access to information each time more complex and specialized. Couzinet brings up the appearance of the first thesauri, that combine expressions and words, alphabetic ordering, permuted and hierarchical. We understand that the fact that, for the author, the structure of the list is able to aggregate different sublists and permutations, and that its structure might always be in negotiation during its use, that all these traits approximate lists from the dynamics of devices for social interaction. In fact, according to Couzinet (2012, p.138-139), lists are info-communicational devices. In this sense, databases substitute stable order by a set of fields rearrangeable according to queries or interfaces for presentation. They are like a superposition of lists and tables, that will become a standard form of knowledge storage and circulation in information technologies.

As we pointed out in the previous chapter, Manovich (2012) talks about the influence of design patterns in the access of information. He also explains us that the patterns that are shared across many devices may become what he calls forms of information. In a previous article (1998), he had already defined the database as one of these forms: he pondered, like Couzinet, that databases are basic structures for informatics, storing lists of objects that can be reordered and recombined. Unlike Couzinet, he adds an interesting remark: the database, for its multidirectional relations, is like a catalog of objects organized in space, rather than time, and in this sense it is opposed to the narrative, another historically dominant way of organizing information. In the article of 2012, he takes this discussion a bit further, and posits that, with social media, there is the emergence of another dominant form, the data stream, in which, instead of browsing a collection of elements, the user would experience a continuous flow of events. He also adds that, although in platforms like Facebook and Twitter the experience seems to be very passive, it is always somehow configured by the user.

We would like to point out that, from our point of view, there is indeed a difference of experience when the interface is like a news feed and when it is organized like and archive, but we should always keep in mind that these two are interchangeable, and can even be based upon the same data. The first emphasizes recombinations while the second emphasizes modulations, that is, not an objective control of what the news feed will show, but configurations based on general criteria of privacy, priority and connections to information

sources (in the case of Twitter, the profiles you follow). Nevertheless, one experience turns into another once, for example, you try to manage your list of friends in Facebook or, on the other hand, you see an archive that is updated in real time as a news feed. In this sense, our lists, tables and grids are not information forms, but information tools that are in the roots of many translations and conversions.

So with personal computers and more intensely with the internet, there is a multiplication of personal and collaborative tools that use lists, tables and grids in their visible displays or organize their inner processes through these same objects. So apart from helping institutions to organize and store specialist knowledge, these structural tools will be part of our common practices, be it in a more evident manner, be it in more sophisticated and subtle applications. Many of those devices that are based on and that display lists, tables and grids will be, in fact, a stage for many disputes between logics of categorization and ordering proposed by specialists, that inscribe their knowledge and priorities in the resources offered by the devices, and other forms that emerge from ordinary practices and narratives.

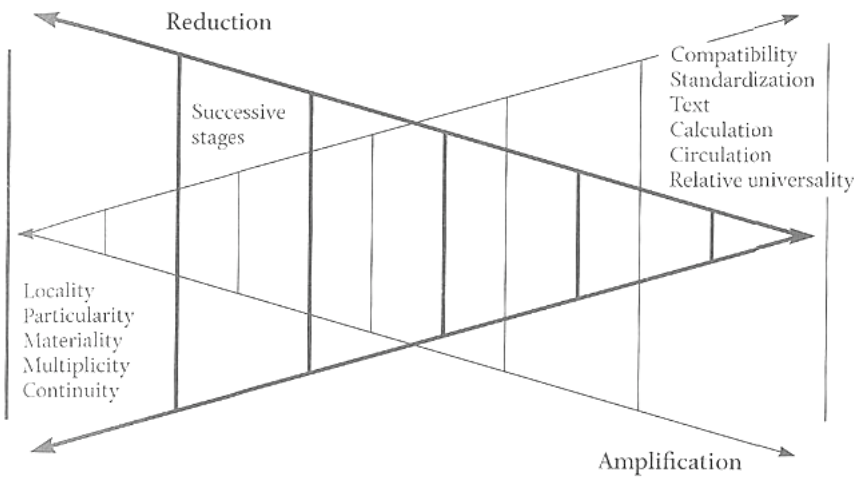


Figure 4.4.1. Chain of information-producing transformations in scientific inquiries. Source: Latour,1999, p.71

Latour, in his beautiful diagram (figure 4.4.1), aligns many transformations where research objects are defined from the complexity of the world and become at each step more compatible, standardized, textual and relatively universal, while carrying through the methodological links with the starting references. Because we are looking at forms of visualization and basic structures like lists, tables and grids, we realize that, in the display of content, there are similar steps, but they range between data and narrative, and they may not be sequential like what we see when looking at the process of research as a whole. Even in the process described by Latour, there is this oscillation of conversions between lists, tables and

grids, back and forth, not necessarily in this order. We want to consider these three objects as filters, or encoding/decoding systems for the many translations that are necessary for the development of visual analyses and demonstrations through information visualization. Even in the decoding stages, newspapers and also scientific publications develop many strategies to guide the reader through many passages, back and forth, between data and narrative, during the process of reading. These echoes the difference between data and literature that for many must be so stressed. Nevertheless, in the context of this study we will keep this in mind as a guideline: while revealing the data structures, look for processes of transformation between data and narrative that develop information and knowledge.

5. FOURTH CHAPTER: VISUALIZATION AND ITS PRACTICES — SOCIAL SCIENCES AND JOURNALISM

As previously stated, for this thesis we were able to conduct fieldwork in two specific fields: data journalism and controversy mapping, each one inserted in the broader fields of journalism and of social sciences, respectively. In this chapter we intend to: characterize data journalism and controversy mapping, describing what differentiates them from other methods and practices in the respective larger fields and outlining aspects about the usage of visualization tools; describe our methodology and tasks carried out in fieldwork; and discuss some preliminary findings. But first it is important to briefly outline some traditions in the broader fields in which data journalism and controversy mapping are inserted.

According to Spence & Wainer (in PLAYFAIR, 2005), even though large collections of statistical data were widely available even by the second half of the seventeenth century, the first major experiments with statistical graphics would only start from the end of the eighteenth century on, with Playfair's Commercial and Political Atlas, addressing macroeconomic data on international trades. This publication was a collection of statistical graphics and text commentaries and was intended as a work of economic theory, but also proposed this new format for displaying hard facts. He believed that it was possible to go without the data tables, since the graphics communicated so much better. Nevertheless, the scientific ethos of the time was characterized by some resistance to images in general as aids for scientific work outside the natural sciences, and critics expressed some concerns with the possibility of misunderstandings and distortions, in comparison to the raw values organized in tables. Adding to this general distrust of the images and the senses, there were the technical difficulties of printing images, that also significantly delayed a broader adoption of graphics. Nevertheless, as we previously discussed, by the end of the nineteenth century, the sociologist Tarde (1883) had already proposed, quite enthusiastically, that statistical graphics would be the social equivalent of the organs of the senses, providing a clear view on current events, and especially bringing back some notion of continuity in the relation between discontinuous records. As we could say nowadays, it would allow the readers to move from the record to the pattern. He compared the trace of lines in line graphs to the flight of a swallow, displaying contextualized intensities instead of fragmented data. He advocated their broad use in newspapers, for the general public.

According to Healy & Moody (2014), towards the end of the nineteenth century and in the beginning of the twentieth century, sociological work was published with rich graphs using varied methods. But, for some time afterwards, graphs became once again very rare in these publications, and quite complex data tables became the rule for displaying quantitative data. Among the many reasons ascribed by the authors is the possibility that, in the quick rise of causal-inferential modeling methods, statistical images were considered like common figures, taken mainly as descriptive resources. There was a concern that the visual display might lead do imprecisions in the interpretation of statistical data. Paradoxically, visualization might have become a victim of early adoption of quantitative methods in social sciences, that valued the regularity of numbers and developed reservations towards visual representations. At the time, the tables, even if they were very difficult to decipher, did manage to contain actual data and not summaries, because of the smaller size of datasets. Later on, even as visualization started to be recognized as a valid resource for analysing data and displaying results, there was a pressure that every graph be supported by data tables, to allow the replication of results and the use of the same material in other works.

Recently some technological changes allowed for a different approach of data visualization in social sciences:

> "The utility of visualization methods — in particular their ability to effectively summarize very large quantities of data or very sophisticated modeling techniques — is partly dependent on related developments in the areas of data sharing and reproducible research. If data is accessible as needed, using figures instead of tables becomes much easier." (HEALY & MOODY, 2014, p.5)

We should note that they are not exactly talking about the technical reproduction of graphics, but about the technical possibilities of making base data available and accessible for verification, replication and reuse, without the necessity of publishing them directly into scientific journals in the form of enormous tables.

Healy & Moody make a distinction between visualization for exploration and for presenting final findings. Visualization for exploration is associated with exploratory data analysis, which involves using graphical tools for rapid interactive visualization and discovery that is openly inductive. These methods are for some regarded as too unstructured, allowing for the risk of facilitating spurious associations. Nevertheless, as the size and complexity of datasets increase, exploratory data analysis becomes more necessary, and at the same time is coupled with new methods for confirmation and validation. These techniques are nowadays in the

center of social cartography like controversy mapping. On the other hand, visualization for presenting results has a rhetorical component, at the very least to convince the audience of one's findings, and would face the challenge of being effective in relation to one's arguments and at the same time faithful to the data.

For Offenhuber (2010), there are also two main purposes for data visualization: thinking and showing, which would be fulfilled respectively by an exploratory function and a rhetorical function. The exploratory function is closely related to data analysis and in information sciences it is linked to objectives such as of amplifying cognition and crystallizing knowledge. In this case, "the story is inside, implicit in the data" (p.367). The rhetorical function, on the other hand, is most theorized by designers and other fields that deal with forms of communication. Visualization, as any inscription, is also the place for narrativity and polemics, it is guided by the possibility of sharing. In this case, the story is the social narrative associated with the visual representation, and it comes from outside, from the discursive context of the visualization. For Offenhuber, visualization should be understood as a representation of meaning. This meaning is neither a preexistent nor an underlying feature of data: it comes from an interpretation of data when it is structured in a specific way. Therefore, visualization sets an ambiguous space where the exploratory and rhetorical functions are intertwined.

Visualization is traditionally used in journalism as an artifice for visual explanation: it makes certain kinds of information clearer. Mario Kano (2013), coordinator of infographics at Folha de São Paulo, discusses the idea of a visual journalism, and explains that some informations are simply better expressed in images. So, it is not a matter of turning every bit of information into graphics, but understanding which bits are clearer if displayed visually. He argues that, historically, newspapers separated image from text, mainly for restrictions in the printing technologies. This created an artificial separation between the work related to producing text and to producing images in newsrooms, and delayed for many decades the broader adoption of statistical or otherwise quantitative graphics in newspapers. With offset and color prints, these two strategies for informing started to be reconciled. Apart from quantitative graphs and schematic drawings, he refers to other sorts of visual arrangements, like bullet lists with icons, that would aid quick and non-linear reading in his conception of visual journalism.

This resonates, we believe, with the discussion we proposed in the last chapter, about a progressive fragmentation of content, as yet another strategy for fragmenting and

reassembling. Also, while working with the broad concept of visual journalism, visualization is framed inside the logic of the visual explanation (TUFTE, 1997), in a narrow sense. However complex and important this function may be, we believe it to be an indication that in many journalistic traditions this rhetorical feature of visualizations that is indicated by Offenhuber tended to get stripped from the exploratory features and from giving access to meanings that might emerge directly from the visual display. There is a story to be told in text and image and published visualizations will mostly point to stable aspects of this narrative, that will work with objects and references that are outside the data. This effect, if it is not coupled with some exploratory dimension, can be very restrictive in terms of the experience with the visualization, making it mostly an appendix, an empty shell to be signified by the textual narrative.

We should point to the possibility that with the effort of clarifying messages in visualization, one might be reifying previous assumptions, not feeding debate. Also, that the structures of visualization themselves, for the fact that they carry many traditions and will of course highlight specific aspects of the data, might indeed develop biases and hide relevant information. Around the world we find examples of misleading visualizations in the news, especially in graphs and charts of the more traditional kinds, like bar, line or pie graphs:[2] these stabilized and well adopted formats are more naturalized and the reader may take them superficially for their general presentation, without noticing many structural biases, identifiable only in the fine print. Nevertheless, we would like to draw attention to another concern, which is related to the mutual influences between mediatic and public agendas, and how unfounded stereotypes present in public conceptions of social reality may be reified in the news if there is not a careful attention to the visual display of data and its interaction to the more general context of debate where it is going to be decoded. Like in journalistic sourcing, pluralism may be a good solution; in this case, to develop many different graphs in order to see and compare many aspects of the data.

Elections are occasions where even newsrooms who have a very literary tradition will turn to data. Brazil is coming from three consecutive presidential mandates of a traditional left-wing party, PT (Workers' Party). The first two were from Luiz Inácio da Silva, and the third was from Dilma Houssef, who has been recently been reelected for more four years, that started in

[2] For a few examples, see http://gizmodo.com/how-to-lie-with-data-visualization-1563576606, and, for a more in depth discussion, https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis.

the beginning of 2015. One of the main characteristics of these three mandates was the investment in projects aimed at the most poor, like a program of conditional transference of income called Bolsa Família. For that, there was a deep concern, among opposing sectors, that these programs might make the beneficiaries dependent and not lead to a durable effect in terms of better income distribution in the country, and that they might aim at gathering votes for following elections.

Brazil's geography shows how serious the issue of income distribution is, because populations are very concentrated in big urban centers and the states that concentrate the country's industries, services and most of the revenue are in the Southeast region, while the poorest are in the Northeast. So logically in the populations of the states in the Northeast there was a bigger proportion of beneficiaries of Bolsa Família.

In the pre-election polls, the states where Houssef was more popular were mostly in the Northeast. Quantitative graphs published in some news outlets confirmed that there was indeed a correlation: the municipalities where there was a bigger proportion of beneficiaries also have shown a bigger proportion of intentions of vote in Houssef. When the results of the second round of the elections came, most of the news outlets gave the results in maps where each state was color-coded depending on the winner. The Northeast was all bright red for Houssef. The general impression: the Northeast elected Houssef. Social networks were flooded with shaming messages towards the northeasterners, and maps proposing that the nation should be separated by a wall in two independent countries, one with the south and other with the north. Most of these fictional frontiers ignored the varying proportions of votes throughout the all states and municipalities, and also the fact that Houssef had the majority of votes in Minas Gerais (MG) and Rio de Janeiro (RJ), states in the richest region of the country, and that Neves, the other candidate, had the majority in the northern states of Rondonia (RO) and Acre (AC).

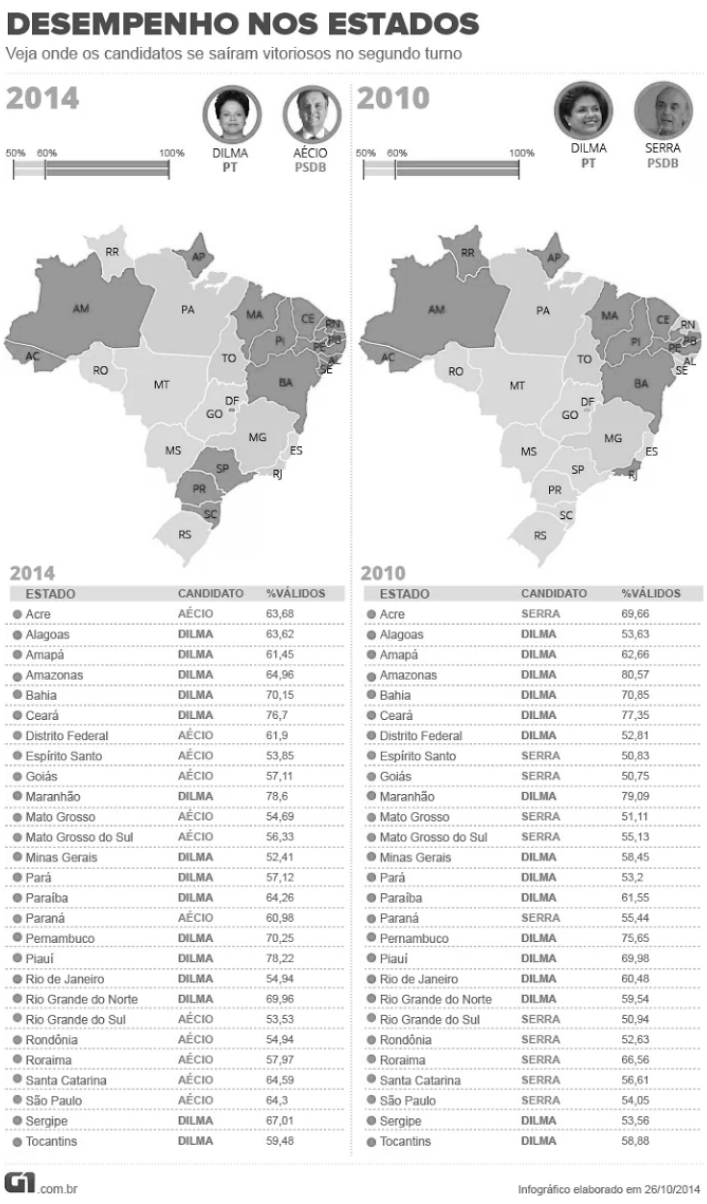


Figure 6.1. Map of the results of the election divided by state with no reference to absolute numbers. It reinforces the concern that the president was elected on the votes of the poorest, who were beneficiaries of programs of conditional income redistribution. Even though the poorest states are in the Northeast, the ones in the Southwest are far more populous, and elections in Brazil are won by absolute numbers (http://www.uol.com.br).

It took a speech from a state deputy for a wake-up call: contrary to the american elections, brazilian elections are not won state by state, but in absolute numbers. Since there was no emphasis on the actual number of votes, and the Northeast of the country has a bigger territory and more states than the Southeast, those maps and even the accompanying tables gave us the wrong impression. If you added the votes for Houssef in the Southeast and compare them with the number of votes in the Northeast, you would find that both regions had given almost the same number of votes towards her election. Even though in all four states of the Southeast region she had only won in two, it is far more populous than the Northeast. After the elections, the news site G1, of the same media group as Jornal O Globo,

published a very rich analysis of the results[3] with many interesting visualizations, but failed to deal with this issue.

Sometimes, the solution for such misinterpretations can be quite simple. The data journalism blog Na Base dos Dados from Jornal O Globo Online, by Gabriela Allegro and Fabio Vasconcellos, has a more experimental approach, and published a bar graph a few days after the election, that gave a more realistic view of the distribution of votes by region. Of course, they realized the misinterpretation that was flooding social media, and decided to show another aspect of the same data set. This shows how the encodings in professional media and the decodings across wider social contexts may interact creating new visual interpretations of data, and how the many versions of visual displays of the same data tell a story of a wider debate. Once again, the use of many visualizations in order to display different aspects of the data, may develop a more realistic representation of social reality and thus feed debate, not reify assumptions.

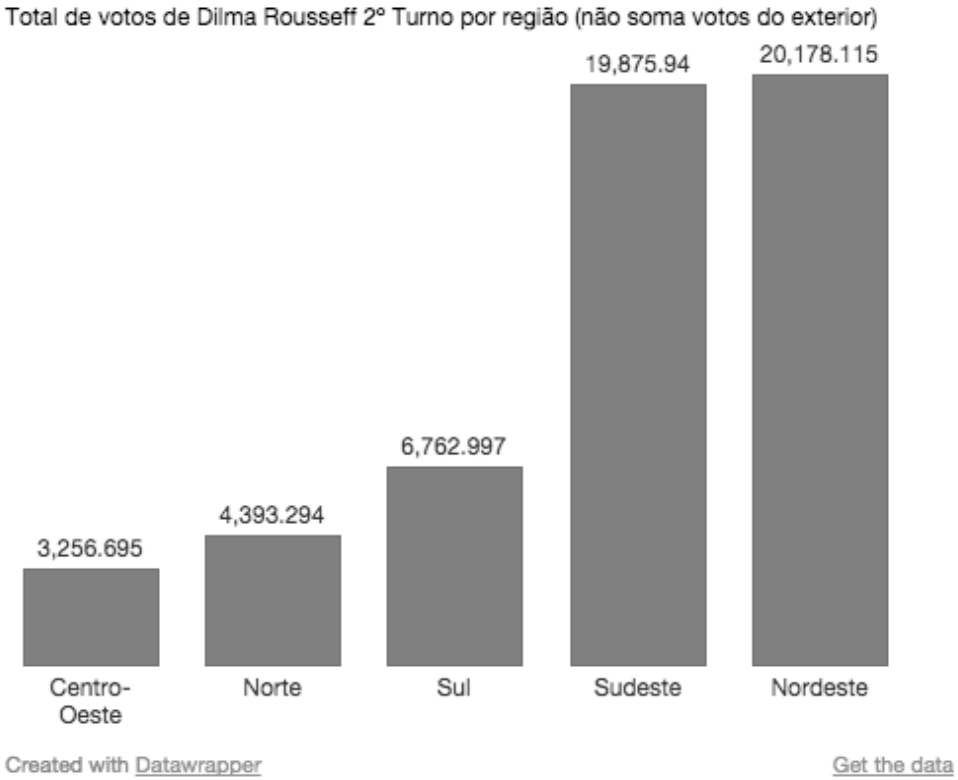


Figure 6.2. Less visually appealing bar graphs display a more realistic view of the distribution of votes by region. Source: http://www.oglobo.globo.com/blogs/base-dados

[3] http://g1.globo.com/politica/eleicoes/2014/blog/eleicao-em-numeros/1.html

Before election, graphs showed the increase of popularity of Houssef in municipalities where there was a higher concentration of Bolsa Família beneficiaries. They were repeatedly used as evidence and crystallized a nebulous concern. The statistical maps of the final results gave the final blow. Even though it is interesting to show the proportion of votes in each state, we believe the maps have been oversimplified and showed only a much restricted aspect of the data. There were many possible graphic solutions to reconcile the raw numbers with the percentages, but the data was presented in an overly aggregated form. So in some cases, visualization can collaborate to reinforce concerns of public agenda, without really providing tools for debate. More than ever, it is important to differentiate issues from ungrounded certainties and stereotypes, that may find in the power of evidence of data and visualization a mistaken confirmation.

The effort of demonstrating clear and well defined facts, which ends up restricting the possibilities for uncovering meaning from data in visualizations, is quite central in traditional newsrooms, and aims to address very real concerns, like avoiding misinterpretations while being well adapted to superficial attention habits. At the same time, in scientific circles, there is a stronger tendency for demonstrating the analysis, and therefore building the narrative from inside the relations displayed in the visualizations used in association with other narratives. Demonstration, in this sense, means reproducing the findings, even if in a staged way, in order to stabilize it as a verifiable scientific fact. This means that the coupling between the exploratory and the rhetorical functions is more effective, as described by Offenhuber. Nevertheless, when a scientist faces the challenge of communicating to a wider public, of people who may not share the references of the narratives of the laboratory life, we see the emergence of discussions that are similar to the ones in journalistic circles, regarding the depictions of social reality: of either developing representations that will stabilize scientific certainties (or social facts), that have no resonance with the controversial and complex reality of scientific work (or social events), or give way to complexity and controversy in displays that may not conquer the same adherence from the part of the public (see, for example, MASSARANI, 1998).

5.1. JOURNALISM, DATA AND VISUALIZATION

5.1.1. What is data journalism

Data journalism emerges in the form of a growing emphasis on data analysis for newsmaking and provides new tools and methods that reposition the practices of traditional journalism, from the newsroom, sourcing and editing to publishing. According to many professionals (GRAY *et alii*, 2012), it is an answer to the growing importance and penetration of data in contemporary societies, which demands that journalists develop new skills in data sourcing and analysis, "to be able to analyze and describe today's events with a discerning eye" (GRAY *et alii*, 2012). We will approach data journalism as a set of practices and methods that can be present throughout the whole news making process and in news publishing, not necessarily replacing traditional methods, but collaborating with them. Data visualization may or may not be present in the final publication, but we realized that, as the practices of data journalism mature and become more robust, visualization tools become important for analysis and more elaborated graphs are published for readers on print and online media.

Of course, data has been used traditionally in journalism: mainly governmental or otherwise institutional data, be it demographic or financial, has been for many decades used to illustrate or describe reported events (TUFTE, 1983). More recently, surveys made by the news outlets themselves show voting intentions for elections and other kinds of popular opinion and demographics. So journalists, depending on the profile of their news outlets, have been for a while more or less comfortable about using and presenting these kinds of data. But nowadays, the web, mobile technology, and, more broadly, the penetration of digital networked technology in several levels of our societies, form a different kind of data landscape where our lives become increasingly and more pervasively described and managed as data. In the one hand, everything one does online leaves traces that can be tracked and even traditional demographics and financial data, due to more refined monitoring methods, become more voluminous and complex. All this data is in general more abundant, accessible and less filtered than traditional data used in social sciences or that were traditionally incorporated in the news, and for that it presents not just a quantitative growth, but in many aspects a qualitative change. So many challenges arise in terms of methods used for its analysis, be it in company strategy, public policy, social research or news reporting. In journalism, this highlights the need for additional skills in data query and analysis, as well as a more critical

approach to data manufacture and its possible biases. It also entails different challenges regarding the ways in which journalists will frame and narrate events.

Journalistic practice has been historically guided by a concern with objectivity, but the ways in which this is translated into actual guidelines and practices has gone through many transformations. In that sense, the promise that data could lead to more reliable reporting mirrors some traditional concerns of journalistic practice that have been on the agenda of professionals for quite a while. Because of that, we understand that data journalism inherits some formulations from previous strands. Meyer (1991) believes that most limitations of modern journalism, like lacking clarity of what are the most important stories to report on, being too dependent on press releases and easily manipulated by many interests, are due to a lack of training in information science or in methods that would help them dealing with the challenges of reporting in a time of information overload. He traces a parallel between journalism and scientific traditions, advocating that both have much in common. He proposes a new strand of journalistic practice, issued from the joining of journalistic traditions and information science, called precision journalism.

Data journalism, in that sense, rather than setting aside objectivity claims, proposes yet another approach to objectivity. Traditional journalism would target objectivity mostly by searching for impartiality and plurality of human and institutional sources for reporting a story and putting these observations into a language that is as clear and faithful as possible. Reporters would trust that the audience would be capable of unpacking the story and critically assess the information that is given towards their own take on events. The methods developed in data journalism would contribute by focusing on analyses that can be verified and replicated again, while trusting data as source and data analysis as a reliable method for describing events, in that sense drawing from contemporary scientific methods. Thus, visualization inside newsrooms may be used much like it is used inside research, that is, to constitute and analyse objects (facts or events, for that matter). But in order to publish visualizations journalists face the challenge of developing new ways of storytelling that would convey this scientific replicability, while still aiming at one of the profession's main duties, which is to fuel debate, to "advocate democracy without advocating particular solutions" (DURHAM, 1998, p.122).

Journalism is one of the main practices that can set the stage for public debate, mediating and translating different social aggregations and forces. Journalists have as their main concern the

goal of reporting for as many citizens as possible, while being as thorough and as unbiased as possible in their sourcing and content. But, in practice, that challenge may represent two opposed needs: for encompassing larger and larger publics, the news tend to be simplified, but, on the other hand, it is hard to be thorough given the speed in which events are spotted nowadays, and due to the complexity of social life and debates in general. In the pages that follow we will try to point out some changes in sourcing and reporting practices since journalism started to incorporate data as one of its sources, and examine some examples of visualization tools that can help meeting these challenges.

5.1.2. Methods and field description

For our empirical work on this issue, we were able to visit some of the main newsrooms in Brazil – Folha de São Paulo, Estado de São Paulo, Jornal O Globo, Globo.com e Jornal O Dia – and to interview about twenty different professionals, among journalists, designers and data specialists that had consistent work in data-based journalism for different kinds of news outlets, inside and outside Brazil. What we came to find were professionals with varied backgrounds and profiles, each one with his/her own very enticing idea of what journalism could advance with data, and how to reveal social issues with data. Our main entry points to this very large discussion was the challenge of changing everyday practices of traditional reporting and how it affected the work and the challenges the newsrooms designers faced.

For data journalism, we interviewed professionals with experience in newsrooms, be it on print or online publications, on daily newspapers or monthly magazines, journalistic associations or design or media firms. They were a total of twenty-five interviewees, among which twelve were designers, ten were journalists and three worked in managing or coordinating positions. Most of them were brazilian, even though some of these worked outside the country. We also had the opportunity of visiting five of the major newsrooms in Brazil, which were Globo.com, Jornal O Globo, Jornal O Dia, Folha de São Paulo and O Estado de São Paulo. Our main goals were: [1] understanding the use of data and visualization in newsrooms from sourcing to publishing: who uses and how? [2] understanding how the use of data and visualization may influence agenda setting, if at all; [3] identify the main concerns of professionals of the field about data and visualization in journalism, and the connection of these resources with public debate through newsmaking. Even if it was not part of our initial

goals, we were also able to gather an interesting outlook on the first major advances of data-driven journalism in Brazil.

With these objectives in mind, we prepared a basic script for the interviews, that could be adjusted according to the specific profile of the professional. The interviews lasted between 45 minutes and one hour and a half, depending on the interviewee's availability and on the discussions raised. Some were presencial, others were conducted through video calls. When necessary, we were able to follow up on the interviews a few days later through e-mail, in order to gather more detail on specific issues.

The general script for the interviews:

Part 1: profile

a) Please tell me a bit about your experience, how it led you into data journalism, and how is your work nowadays.

b) What about the places you work at now and worked at before, how do you think they support the use of data in journalism?

c) How do you see the use of data in journalism nowadays? Are there any fundamental differences when data analysis is used in newsmaking?

d) There are some methods of social research being developed nowadays, which are largely based on data analysis for defining and developing inquiries. How would you compare the scientific approach to the journalistic approach towards data? Do you see any possible approximations or clear frontiers between them?

Part 2: newsmaking

e) How is your work method? From the data: what kinds of data and data sources do you work with; to the investigation itself: what kinds of technologies and methods do you use?

f) Do you work in direct collaboration with other professionals, such as developers, designers or journalists with different specialties? How do you feel about your collaboration?

g) Telling news stories becomes significantly different when they are based on data? To which extent do you feel it is important to explain your methodology and objectively what you found to the common reader?

h) Do you think a story that is based in data can be harder to report? Why?

Part 3: visualization

i) The field of visualization is filled with many terms that overlap each other. How would you trace the difference between information visualization, data visualization, infographics, graphics etc?

j) Do you use visualization to analyse data?

k) What about the final news piece? Do you believe visualization can be useful for telling the story once it is based on data? What challenges does it entail?

l) Do you think the visualizations used in analysis are different from the ones that get published? In what way? Is there a routine for transforming one kind into another?

m) What do you think about the use of visualization in journalism? Do you see any fundamental differences between print, video and web supports in that sense?

n) (Only for outlets that have both print and digital versions) Do you produce visualizations first for digital or for print? Do you have a routine to translate one into the other? What are the challenges in your opinion?

o) How do you see the relation between text and visualization in the published material?

p) Have you ever had any experience working with collaborative or real time data? What do you think about its possibilities and challenges?

q) Do you see any possible uses of visualization in tools for journalists inside newsrooms? Do you have anything of the sort in your company/institution? For example, presenting a dataset on all the available electoral results, so that at every new election they can access the data and look for patterns.

The field of data journalism

We can say that, nowadays, most large news outlets practice data journalism at some point, especially considering that data tables have appeared now and then in the news since the nineteenth century, and that throughout the history of newsmaking the presence and the role of data in western societies just grew, leading, conversely, to the use of many formats of data

display and some methods data analysis. Nevertheless, there are companies and institutions that integrate the use of data in a broader, more structured way. We would like to draft two main experiences that we could identify: first, the intensification of the use of data in reporting can happen through its use in data-driven applications and elaborated visualizations. Outlets that have this profile will tend to have big infographics departments, that put together designers, developers and editors to produce specific features based on data. These can generate news stories or be related with many of them. In these cases, the infographics department has more autonomy than in traditional newsrooms, like it is the case in The New York Times, one of the strongest examples of the use of data visualization in the news in the world. Another experience can come through the creation of specific data journalism departments, that will try to furnish the newsroom with the resources for producing data-driven news. In the case of International Consortium of Investigative Journalists (ICIJ), for example, they created a specific division for data-driven journalism, but, from our point of view, there is a certain continuity between data analysis and some classic perspectives of investigative journalism, that entail checking records and look for traces outside the official discourse. Data analysis comes to equip and deepen investigative activities that were already in place. ICIJ has become a reference in data-driven journalism since they dealt with a leak of offshore bank accounts, in The Offshore Leaks Project. Also, many news outlets will have data specialists in key positions in the newsroom, so that they can give support to the use of data. Other outlets who have relevant work in the field, either using visualization or not, are The Guardian, Le Monde and the ones focused on economics, like The Financial Times.

Even if our goal with this thesis is not to examine the perspectives for data journalism in Brazil particularly, we had the unique opportunity, while visiting some brazilian newsrooms, to observe an interesting moment, where the use of data analysis starts being heavily discussed and tested in newsrooms. The fact is that, traditionally, the use of data in Brazil has been mostly limited to public surveys, and governmental demographical statistics. The contact the journalists had with this data was mostly indirect, meaning they did not actually put their hands on the data, they mainly looked at graphs, tables and analyses that were published by other institutions and transferred it into the news. So it was easy for a research institute to influence this kind of reporting, depending on how they described what the data had to show.

But with digital media the sources of data multiplied, and data became more available, but less aggregated. Many of those news organizations were also conducting editorial projects on the web, and, as the information technologies advanced and became more accessible, these

same projects turned into new kinds of data to be assembled to generate content. Since Brazil has a huge territory and many isolated regions, there is an important governmental effort to use online information systems: the annual revenue tax declarations are done online, and so is the control of taxes in ports all along the brazilian coast, for example. There are many laws for transparency being passed, so many datasets on governmental activities and public budget are now available online. Data became more varied than demographics, but, on the other hand, much more complex.

We started the brazilian part of our fieldwork expecting to find just a minor concern for data among journalists and maybe some deeper reflexion about the issue among designers, since, you know, they built the graphs. This last idea, about the designers, did stand, because, even if data analysis in not common procedure in most brazilian newsrooms, they are the ones who must deliver the graphics and that will work directly with the data. Because of that, they will tend to be the ones who will spot any problems in the data and will have to face the responsibility of finding reasonable workarounds. So the basic issues of working with data are more evident in their daily work, and they tend to have many thoughts to share. Apart from that, we were quite surprised to identify many strategies for dealing with data, some recent, some fully incorporated, and a strong disposition on the part of the professionals for discussing the issue.

The newspaper Folha de São Paulo is known for its use of data-based infographics, that have become their tradition for decades. They have a very strong team working on visualization and many features that have won many prizes under the coordination of Mario Kanno. The Folha de São Paulo Group has a company called DataFolha, for public surveys: it is an independent company, not a division of Folha de São Paulo, but it produces many surveys and demographics to the newspaper and lends likewise service to many other institutions. Even though DataFolha may give the newspaper Folha de São Paulo much autonomy from some of the biases of government demographics, and even considering their outstanding work in visualization, there is still some work to be done so that data analysis can really penetrate the newsroom and reporting practices in general. According to Marcelo Soares da Silva, who works as data specialist in the paper, it is very hard to include data analysis in the routine of the newsroom, because it demands a very different profile of journalist, where most of them are still issued from the point of view a literary tradition. Silva is a journalist, but his interest in data analysis has led him to work with metrics and monitoring of the digital content of the company.

By the time the interviews were conducted, the most known effort to bring data analysis into the regular practices of journalists was Estadão Dados, the data division of Estado de São Paulo. They have assembled a dedicated team of professionals whose the task would be to collect, process and interpret relevant data for either producing data-driven news themselves or equipping the whole newsroom with data-driven applications. At the time of our research, they were a team of four reporters that had mixed training in journalism and development or data analysis, and a senior reporter with a career in journalism of politics and economics. They were quite independent from the rest of the newsroom, a situation that presented some advantages and disadvantages. On the one hand, they were quite free to develop projects and careful analysis and post them on their blog[4], commenting and deepening the approach on the current hard news published on the print and online versions of the paper. They also develop work for the main pages of Estado de São Paulo. On the other hand, they had a hard time getting their point across the newsroom itself. The senior reporter tells us that there are some crude misconceptions that have already been clearly exposed through data analysis, but reporters keep repeating them over and over in their texts. Also, that reporters outside Estadão Dados will seldom look closely at data when it is at hand or look for data sources behind the story they are working on, let alone come for them for support. On the other hand, since they do not have dedicated designers, they have to compete with the rest of the newsroom for the time of the professionals in the infographics division. Most of the visualizations presented in their works are built in automatic tools, based from their more elaborated work of data-driven journalism. When they heard that at GloboEsporte.com, another news outlet from another media group, they had a team that mixed designers, video makers and developers, they were quite amazed and maybe even a bit jealous.

GloboEsporte.com is the sports division of Globo.com, the news portal of Organizações Globo. It is independent of the print Jornal O Globo, that has its own portal, O Globo Online. Indeed, GloboEsporte.com has a very interesting team of designers, but they work from a different perspective than EstadãoDados. While the later is a division composed by journalists and works mostly analysing state policy, the economy and politics in general, at GloboEsporte.com, because they report on sports, there is a different approach. In fact, they have two divisions that deal with data: the infographics division and the FutDados division. The projects developed in the infographics division can vary from the more playful aspect of

[4] http://blog.estadaodados.com

the interactive animations, to the analysis of sports statistics. They have much autonomy to produce independent features that may become prominent in the pages since this is an online publication. On the other hand, the division called FutDados is composed of a team of eight reporters mostly dedicated to reporting soccer through data. The name is a contraction of the words soccer and data, in portuguese. As EstadãoDados, they do not have designers.

The infographics division produces projects that range from classic infographics like animated cartoons, to visual stories of the snowfall variety, that combine images, videos and many forms of text, and data visualizations. They have a team of designers of mixed profiles, from illustration and video-making, to web design and front-end development, to information design. They have some autonomy to produce special features that might not be originated from traditional reporting, like some specials developed for the Fifa World Cup in 2014, but also provide graphics and smaller features for regular news pieces on request. For that, the team also had a special aid, that is a journalist who is in charge of orienting reporters on their demands of infographic material, either refining raw ideas or turning them down. This is an interesting initiative for a publication that has a lot of demand for visual impact, but also aims for relevant content. During the World Cup, they managed to test some new formats and languages, and came out with many projects that were strongly data-driven.

FutDados had an outsourced service that delivered statistics about each match in brazilian state and national championships, while three of those journalists would recheck the data and make different tests for analyzing the events. Apart from the huge importance the sport has in brazilian culture, that heightens the public interest for every single detail, Flavio Campos (personal communication, July 24th, 2014), one of those very specialized data journalists, noted the fact that soccer is very different from basketball or baseball, for example. Those games offer many evident metrics: they have more points in the score, more faults, differences in punctuation… All of those configure metrics that are easily recorded and give you more material for statistical analysis. With soccer, a single goal may define the final match of a championship, so it is common that journalists find themselves wondering for reliable ways to assert whether if the result really represented the quality of both teams in the field. They give support to all the reporters in GloboEsporte.com's newsroom, but their interaction with the infographics division was not very frequent.

There are also the cases like Superinteressante, a monthly print magazine from Editora Abril, dedicated to scientific curiosities and to the popularization of science. They have a very

elaborated work on infographics that has become a reference. In their case, the visualizations are the main part of the news piece, where the text often comes in the form of an introduction and of annotations and more detailed comments, that elaborate on the visualizations. Often, visualizations are produced at the same time for data exploration and final publication, and the news piece is produced in a close collaboration between designer and journalist.

According to designers and also journalists who already have a profile turned towards data-driven news, like Silva, the data specialist of Folha de São Paulo, Allegro, the data journalist of Jornal O Globo and the crew from Estadão Dados, the biggest challenge for implementing data-driven news in their newsrooms is inserting the culture of using data into the everyday work of the common reporter. In general terms, they point out that reporting with data demands at least using tools like spreadsheets and acquiring some basic knowledge of data analysis. They also understand that the biggest challenge for these reporters is counterbalancing a routine of sourcing with people and using declarations with the realization that behind most news stories there is data to be explored, in order to ground reporting and improve the story.

Paiva, chief of infographics of Jornal O Globo, points to another problem, which is the classic separation between journalists and infographics or art departments in most traditional newsrooms: to an extreme, designers are seen as technicians or decorators whose work does not collaborate for producing relevant content. This goes back to older discussions in the field about visual narratives (KANNO, 2013), that also point to the need for closer partnerships and interaction between journalists and designers, with a bit more autonomy for these last. As digital and interactive content grows and many newsrooms turn to the “digital first” approach, this discussion takes on a new aspect. Designers need to acquire specific knowledge for developing complex projects that demand a point of view of data analysis. Because of that, in many cases, they are going to have a clearer understanding of the impact of a dataset to the news piece and could take on an important role in producing data-driven content. To profit from this, it is necessary to find new paths or alternate routines in everyday work and, in an extreme, considering the possibility of reorganizing the structure of newsrooms, so that more interaction between teams can be reached and other kinds of professionals can be integrated, like specialized developers and data analysts.

So in most of the newsrooms we visited, even though there was not a mature solution of integrating data analysis in reporting, there was some concern about the use of data, about

how it could ground the reporter's claims, about ways in which it could be integrated in regular reporting, orient traditional sourcing and reveal other aspects of events. The use of data also seems to give rise many issues on team composition and on the interaction between different teams and professional profiles. The questions of how to integrate developers in newsrooms, of whether to understand designers also as analysts and content producers might be very challenging for traditional newsrooms that tend to pull apart content and image, leading to frozen structures and less collaboration between different specialties. Also, there is the issue of how data should be displayed, how it could reach more people, and whether if source data should also be published, as a strategy for participation. And things are evolving quite fast: few days after we had the chance to talk to Gabriela Allegro, data journalist of Jornal O Globo, we heard the news that they were starting their own data division, with her on the lead. At the same time, she started a blog called *Na Base dos Dados*, with another data journalist called Fabio Vasconcellos. They use it to experiment with data-driven analyses, and develop ways of publishing them as part of the regular content of Jornal O Globo.

5.1.3. Discussion

Before commencing fieldwork, we mostly believed the work with data, and especially big data, had an association with scientific methods and more reliable metrics, and therefore could help avoiding some biases in newsmaking. Also, that by publishing more visualizations, more dimensions of data could be available for the reader, allowing for more interpretation and a critical approach.

One of the first surprises we had in the field, as brilliantly summarized by Marcelo Soares da Silva, data specialist from Folha de São Paulo (personal communication, August 11th, 2014), was that "data is just like any other interviewee": if you torture data, it will tell you whatever you want to hear. In fact, we found that this expression "torture the data" is very current in the newsrooms we visited, always with a negative connotation. So it is not as if data itself could bring more credibility to newsmaking, but maybe integrating more reliable methods of data analysis could have a positive impact on all aspects of the practice. In newsrooms that have a more literary tradition, like most in Brazil, even for news pieces that use statistics or economic indexes, data comes already packaged by statistical institutions through aggregated reports or translated by specialists' declarations. So it is as if data was used indirectly because, strictly saying, it is not actually analysed in the newsroom.

What we found was that, even though there is a strong discussion circulating newsrooms about regaining credibility to journalism through the use of data, the most important aspect of integrating data analysis into newsrooms seems to be that it forces reporters and editors to reassess weak points in the whole process of newsmaking. According to Cabra (personal communication, august 25th, 2014), coordinator of the Data Division at ICIJ, investigative journalists are used to looking at records and finding details and discrepancies. But, in order to take a step back, examine the wider set of evidences and see the patterns (especially in today's huge records) you need a data analysis approach: "The reason I fell in love with data is because it allows me to go from the *case* to the *pattern*." Journalists will use data to find and frame issues for the news pieces, while building a clearer perception of their relevance in a wider context. Of course, this has deep effects on agenda setting. Moving on, data is very seldom the only source for reporting: the *owners of the data*, people who gathered or generated the dataset, are important sources, because they will give precious information about its underlying criteria, original purposes and possible biases. This maps out what can and what can't be inferred or extrapolated from a dataset.

So here we have an example of a practice that addresses the need for a critical approach towards data. Also, once the main actors and issues are framed, this opens a larger field of sources of information to investigate outside the computer. It is always important to translate issues in terms of human experience, so interviewing people who have been affected by or witnessed an event or issue remains crucial. Before data, finding these people could be quite random, and the reporter would tend to rely on good quotes as evidence to build the story. "Data takes out part of that randomness" (personal communication, august 25th, 2014), serving as a road map to decide who are the relevant sources.

We could say that if the starting point for a news piece is a dataset, less trust has to rest on institutions and specialists as reliable sources, and more effort must be done in terms of analysing the data while understanding its initial purposes and methods and avoiding its specific biases. We could also say that data makes a better source than institutional discourse represented in the person of specialist, a manager or a politician, that have the tendency of representing interests, propagating power structures and concentrations of social power and not representing issues from the point of view of the affected people. But, on the other hand, we have to consider that avoiding biases and correcting possible distortions in data entails more than developing reliable methods of data analysis. It demands re-connecting data to those institutions of origin. To an extreme, data must be considered as a discourse that can

also propagate these same interests and concentrations. So data and traditional sources present different kinds of biases, but working with both helps developing a critical standpoint that deeply collaborates for the relevance and quality of the news. What we find especially interesting about Cabra's point of view is that it integrates both perspectives and sets a working environment where they can check one another for more reliability.

As for the published content, an important point that seems to be emerging among our interviewees is of always giving account for the sources of data and methods used to treat them and get to the story. Usually they believe it is paramount to give access to base data as much as possible, so that other actors can reproduce results and bring new insight to the table. Many believe this transparency of procedures collaborates for credibility. Allegro, coordinator of the new data division at Jornal O Globo also advocates for maximum transparency regarding datasets, sources and methods used. We asked Allegro about the tendency of many media outlets to create strategies to guarantee exclusive access to sources and how it could be adapted to data-driven journalism. She just smiled and used a very brazilian expression, that could be roughly translated to something like "This is gone." It is not up to media outlets to control the channels of information, this is (fortunately) gone. It is better to accept that and work from there, sharing and encouraging exchange. According to Bramatti, reporter of Estadão Dados, they also do not see the use in keeping raw information to themselves: journalistic practice derives its value in analysis and contextualization. In Estado de São Paulo newspaper, some visualizations are produced for analysis and published in the Estadão Dados' blog. They may serve as basic reference for reporters about specific subjects, thus supporting the production of many news pieces, but at the same time are accessible to interested readers and even to other news outlets.

Visualization in data journalism

Mario leite, coordinator of the infographics division of GloboEsporte.com (the sports section of Globo.com news portal) says that he would rather have his division called Visual Narratives Division. He explains that this is because narratives are what they do out of scattered things like data, images from many sources, illustrations, videos and so on: they create content to be visually explored, that develops narratives through exploration.

There is, indeed, much confusion in the terminology used for visualization inside news outlets. Old school journalists in Brazil will call "art" every image that is not photographic.

This is already an outdated approach, mostly from a time when newspapers used images to fill in blank space or as decorations. Then there are infographics, the most current term nowadays: it refers to graphic displays both of data or information, and to visual explanations in general. So a quantitative graph of the inflation and a drawing that would show the insides of a nuclear submarine are both infographics. They are usually used as visual explanations that will complement the text.

Like previously discussed, visualizations (or graphs) have been used in the news for quite some time now, but, the same way data was not in general an integrated part of the process of newsmaking, neither was data visualization. As the designer Gianordoli points out (personal communication, June 27th, 2014), if you do not have data as part of the process of construction of the news, the visualization will not be something relevant, it will necessarily be an illustration. Indeed, graphs become more like visual explanations (TUFTE, 1997) of hardened facts (LATOUR, 1985), which constitutes a very powerful function and a difficult task, but is still a sign that data is being used indirectly. When it is possible to integrate a deeper data analysis approach, many visualizations may be tried out so interesting patterns may emerge, and then other journalistic methods might be incorporated to develop the story, like described by Cabra. The final piece may or may not contain visualizations, but chances are that the ones that do get published will be much different from the ones used in analysis: apart from being more polished visually, they will demonstrate the most relevant aspects of data where patterns were found. They will accompany the text mostly in a complementary way, but, apart from explaining hardened facts, they may also be capable of demonstrating part of the discussion itself, since the process included visualizing. So visualizations incorporate a relevant part of the news story being told, and become one more resource so that the published story may account for its own assemblage. We believe that this change of focus for visualizations highlights the controversial aspects of every news story and might prompt debate through visualizing.

On a related aspect, as the amount and complexity of the data used in newsrooms starts growing, there is also a search for richer solutions, many times interactive, for the visualization of data, and a deeper discussion about its uses and resources. Also, once the perspective of the media professionals about the control of information and information channels start changing, they start incorporating auxiliary spaces to the main outlets. One example is the blog of Estadão Dados, that offers visualizations with full datasets, not aimed at specific news pieces, and the blog Na Base dos Dados, from jornal O Globo, where Allegro

and Vasconcellos publish more experimental data-based content. In these spaces, digital networks help reorganizing some flows: this might lead to a more fluid exchange between the spaces of analysis and the spaces of explanation.

But all these experimentations do bring many concerns regarding the reception by the ones who actually read the news piece: of course, hardened facts are much easier to understand and adopt on a daily basis, while understanding processes and insights may demand too much dedication for most people. Will the readers understand what is being shown, let alone be interested in delving in such discussions through exploring sometimes complex and unusual visualizations? Carlos Lemos, designer at GloboEsporte.com says that he does not work for an imaginary average reader, but to create a dialogic relationship (personal communication, May 28th, 2014). This comes from an understanding of the fact that each visualization configures a specific context for reader-medium exchange. On the other hand, Marcio Leite, coordinator of infographics also at GloboEsporte.com, believes that this old fashioned abstraction of the average reader is alive and well: from experience, he has learned to respect the limits to which some people will cling in order to understand a visualization (personal communication, July 24th, 2014). So, the larger and more varied is the audience, the more didactic we tend to be. While Lemos points to a clearly explorative perspective that will assemble a narrative along the way, Leite points to simplified visualizations and sequential tours and other resources that will take the user by the hand. These are examples of two conflictual perspectives that generate many interesting discussions, especially in the fields of communication and design. Apart from different standpoints regarding design practices, there are also the traditions of journalistic editing at play: according to our interviewees, editors will usually prefer to present a finished story, and that usually does not get along very well with a more exploratory perspective which might present conflictual aspects of the story.

From what we could observe, the traditional and practical answer to this discussion is simplification: most believe the reader of a daily newspaper, for example, does not have time to deal with complex visual narratives, let alone explore interactive ones that force them to learn yet another set of controls, another user interface. Nevertheless, we could identify a few experiments for achieving a more productive experience with visualization, either in print or online publications. To put it briefly, we believe that this didactic effort should not be denied or accepted entirely: if being didactic means giving broader view of the process of constitution of a news story, it could be very interesting for instrumenting debate and composing shared landscapes. If it means assuming that the public needs always simplified

content in order to be able follow it, then we might just be under the impression that didactics works towards obscuring complexity, and not towards giving access to many possible understandings of it.

5.2. DIGITAL METHODS OF SOCIAL RESEARCH AND CONTROVERSY MAPPING

5.2.1. What is controversy mapping

Controversy mapping is a research method based largely on actor-network theory (TAR), and bred in an intersection between the general fields of Digital Humanities and Science Studies. The first emerged especially from the 1980s on from humanities computing, and nowadays became a research practice that combines methodologies from the humanities with computing tools to analyse both digitized and digitally native (ROGERS, 2011) materials. Inquiries in the Digital Humanities may encompass concerns from humanistic fields like history, literature, medieval studies and the social sciences, and use tools like data mining, statistics, text mining and data visualization. An example of such inquiries would be, for example, seeking, in a literary corpus, patterns of word occurrence that may point to authorship claims or networks of influence. Of course, this promising approach generates quite a few challenges for equalizing the largely quantitative approach of computational methods with the critical and qualitative traditions of humanistic fields. On the other hand, Science Studies, or as it came to be called, Science, Technology and Society (STS), is a field that studies scientific activity from the point of view of the broader sociological, philosophical and historical contexts. It is also highly interdisciplinary, including inquiries that discuss the production of scientific knowledge as a social process, and its representation and interpretation by sectors of society and society as a whole (LATOUR, 1999). Variations like the Scientific Humanities will address the role of scientific controversies in wider social contexts.

Actor-network theory (ANT) is developed initially inside STS research, from the work of researchers like Michel Callon, Bruno Latour and John Law, and later becomes wider, moving on to address the many areas of interest of sociology and related fields. According to Law (1992), starting from a discussion about the production of scientific knowledge as a social process, the funding authors of ANT will highlight the importance of other actors besides human actors, like instruments, practices, inscriptions and a host of other objects that

are interconnected, interact and transform themselves to form the heterogeneous network that society is. The central idea is that society itself would not be possible if it were not for all this heterogeneity: after all, objects are the instances that reify and give permanence to social relations (LATOUR, 2009). Indeed, ANT understands that what makes up the social are groups and aggregates (networks) of actors and their relations. Things such as institutions and knowledge systems are mostly instruments we use to perpetuate or continue certain sets of social relations (LATOUR, 2009; LAW, 1992). Perhaps two of the most powerful points of ANT are, firstly, the idea that the category of actors is not restricted to humans, meaning that objects (including semiotic ones) also have agency, that is, carry intention and gather and translate actions and intentions. Secondly, that there is nothing under, behind or above actors' actions and the networks that they form, no superior law that would explain and guide their interactions, but interactions themselves that get more or less crystallized in the form of aggregates, institutions, practices etc. Everything that merits investigation must come from the actual traces of those relations, so, in order to study the social, we should simply follow the actors' traces. Therefore, the task of sociology will be to characterize the networks of the social it all their heterogeneity, and examine how they could generate effects like the structuring of institutions, the creation of inequalities and the concentration of social power (LAW, 1992).

To many researchers, ANT cannot be considered a full theory: it should be taken more like a set of methodological indications to the social sciences, that seek to avoid that preexisting theoretical frameworks reproduce and naturalize the power structures and generalizing theoretical assumptions. Venturini (personal communication, August 11th, 2014) explains that, while trying to explain society, or that which is social, many branches of social theory tend to separate it in two levels: the level of structure and the level of local interactions. So they end up locating their efforts in understanding how structures configure interactions, or how we use structures in our interactions. Then, ANT comes as a set of recommendations on how *not* to look at social phenomena: it refuses this binary separation and postulates that society is flat, and all actors work on the same level, so we should always begin social research and analysis by following the social actor’s traces, the actual traces of local interactions, before trying to impose preformatted structures and patterns. And, because this is mainly a negative argument, it can be difficult to be taught and put into practice. The visible traces of social actors should be the first and foremost material for the social sciences. Controversies are regarded, in that sense, as privileged environments for ANT research,

because inside disputes, the contents and crystallized relations, that tend to otherwise fade to the background of habit and tacit assumptions, are exposed and become traceable.

Controversy mapping is largely based on ANT: it began as a method for demonstrating its principles. Even though the ANT could be applied to research using paper and pencil and not necessarily with digital tools, controversy mapping has established itself mainly in the digital environment, as a set of methods for mapping and tracking controversial issues from traces left by actors in digital networks, through communication and information technologies (LATOUR *et alii*, 2012; VENTURINI, 2009; VENTURINI *et alii*, 2010). As the experience with digital methods advanced, controversy mapping matured and became something bigger, more like a full-fledged method that would use data produced inside communication and information technologies to map social interactions and analyse social issues in general (Venturini, personal communication, October, 9th, 2014). The data used can come from many different sources, either governmental, institutional, from other researches or from media outlets and social media: the fact that more aspects of social life are being digitized and tracked by information technologies opens up the possibility of tracking social phenomena in the making, in their controversial aspects, and not only when they get stabilized into things like laws, documents and institutions. This represents an important transition from traditional social research, also because it understands online data as a source for social analysis in general, not restricted to online culture (ROGERS, 2009).

In terms of practice, controversy mapping may start with a list: a controversial issue is chosen, and one should just list the statements involved. With that, one can track the debates or discussions involved, and determine the actors that take part in it. Then it is possible to move on to assemble networks that display proximities and distances, as well as alliances and disagreements between actors, and to develop a fuller understanding about the controversy and about social interactions (VENTURINI *et alii*, 2012). Many different sources may be used, along with many tools and methods for extracting, treating and visualizing data. As controversy mapping comes from a deep concern for social interactions, the network maps and social network analysis will tend to be central in cartographic work, but they are not alone, as many other aspects of data will need to be explored, and therefore will demand other tools and kinds of visualization. The finished results will be published in the form of an atlas, with many different visualizations generated for various datasets that may show different aspects of the controversy, and a narrative, a scientific text that will expose methods, source data and discuss the maps and findings. Controversy mapping scholars at the médialab at

Sciences Po have been devoting efforts towards opening the cartographic process to the participation of engaged actors of the controversies and for returning these atlases as tools for public debate and citizenship.

Controversy mapping, as a method of digital social research, aims to explore the possibilities of opportunities derived of the massive amount of data about social interactions that is available today, mostly on the web. For Rogers (2009, 2011), digital research methods inside the social sciences should work on advancing digitally native methods that would go beyond the use of digital tools to amplify traditional ones. So, instead of celebrating that, for example, traditional tools such as questionnaires can now reach a higher number of respondents and spread to wider geographical areas through online forms, different methods based on traces left by networked communications and monitoring start to be tested and refined. Some of them come from information technology environments, but take on different approaches derived from social sciences' standpoints and concerns. Overall, the main idea is that, instead of using digital technologies to further traditional methods or either studying how digital networks modify the ways in which we communicate and organize our society's networks (which would take us to media studies or media culture studies), we should be focusing on the data produced inside the medium as a source for broader evidence and analysis of issues of society in general: "virtual methods and user studies in the social sciences and the humanities have shifted the attention away from the data of the medium and the opportunities for study far more than online culture." (ROGERS, 2009, p.2)

This discussion relates to what Venturini & Latour (2010b) describe as an artificial qualitative/quantitative divide in the social sciences. According to them, natural sciences are in a much better position to track the phenomena they study: technologies and devices like particle accelerators or petri dishes and microscopes allow scientists to track natural phenomena in the making. Meanwhile, the local interactions that compose the social fabric are much more difficult to follow. Objects like written records, laws or institutions, often the raw materials of the social sciences, are the result of much work done by actors in many interactions, so they display the stabilized results, not the processes by which social forms emerge and are reinforced.

According to the authors, in the early days of Sociology, statistics came as a revolutionary tool, because, by calculating averages, probabilities and estimates, social scientists could now describe social structures, project and intervene. Nevertheless, this created an artificial

separation between the micro level of the interactions and the macro level of the structures, and the demand for formulae that would explain how one level influenced the other. Also, a separation between the micro level, that could be studied by means of qualitative methods of research, and the macro level, that demanded a quantitative approach.

In societies that are heavily mediated by information and communication technologies, we come to a culture of the commentary (JOHNSON, 2001), where the conversation that was always present between different works, in the form of influences, references and appropriations, becomes more visible but, at the same time, there is a hyperproduction of documents, and those tend to be less compliant to disciplinary knowledge and classic hierarchized categorization. Therefore, there is a growth in the demand, not only for filters, but especially for instruments that might mediate the access to available data and might work as interfaces to organize information. Answers to this concern are developed in fields such as information sciences, human-computer interaction, communication and media studies, among others.

According to Venturini & Latour, this new affluence of digital traces presents itself as an unparalleled opportunity for the social sciences, because it offers the possibility of tracking, representing and analysing social reality from local interactions, and to render visible the work of assemblage and maintenance of otherwise seemingly stable social structures. Quali-quantitative methods emerge as an attempt to bridge the qualitative and quantitative divide: far from being just a juxtaposition between statistical analysis and ethnographic observation, they propose procedures to track each of the actors and interactions in social phenomena, something that was impossible before the existence and wide distribution of digital technologies.

But this approach entails many challenges. For Venturini (2012b), we live contradictory times, because all this richness in data is not produced inside social research, nor according to its criteria: social sciences have done nothing to deserve such abundance, and now face the challenge of adapting this “second hand data” (p.2) for the interests of scientific research. Part of this challenge comes actually from the level of detail, granularity and acceleration of their production inside the networks of information, so that it becomes difficult to define an interesting outline for a controversy at hand and even more complicated to have a relevant understanding of its movements. But this is the kind of positive challenge, because it is very productive, as we might see from the examples to be discussed: it points to the possibility of

following a controversy in detail and of attaining a less static description of it, unfolding movements in time from the traces of digital mediations.

On the other hand, there is also the concern that controversy mapping in the social sciences, while using this second-hand data, might be inadvertently studying the effects of media structures, rather than studying social phenomena. This is because data that is extracted from social media, for example, carries traits of the platform. In Twitter, for example, we have hashtags that categorise and structure the content horizontally. We should always question how these categories that were created inside communication systems should be adapted to the investigation at hand. According to what Marres & Gerlitz propose, while observing mapping results: "proportional measures (frequency) are more likely to direct our attention to medium-specific dynamics (busting; hyping), while relational measures (connectedness) can help to foreground some substantive dynamics." (2014, p.19). This is one of the reasons why network analysis is so strong in controversy mapping: it allows scholars to develop a sensibility for the relations between actors and their density (Venturini, personal communication, 2014).

There is, according to Rogers (2009), the need for what he calls a *grounding* of the digital traces so that scientific findings really refer to social interactions and not to effects of platform. Controversy mapping may then be associated to ethnographic, anthropological or geographical tools, or from other traditions, that may contribute for a critical position towards the data. Besides that, it also becomes clear that, as a complement to more careful methods for extracting, treating and visualizing data, it is necessary to develop a critical understanding of the workings of the medium that originated the data in the first place. We believe this entails a return to the communication theories, that engage many theoretical resources that are relevant to the discussion, like semiotics, discourse analysis, cultural critique and aesthetics.

We should note how controversy mapping also derives its collective aspect from a fundamental transit between discourse, argumentative positions, meaning attributions and people: semiotic objects, like terms or keywords, are the starting source for assembling a landscape, but they will organize statements, which will lead to outlining debates, and to identifying actors and networks (VENTURINI *et alii*, 2012). This interchangeability and these transformations along entities, described as variations in complexity, are very much aligned with the general principles of ANT, in which the method of controversy mapping is based: social reality and its controversies are composed by networks of many kinds of actors,

being them people and institutions, objects or even semiotic/discursive objects like terms, keywords or statements. One of the most innovative aspect of ANT is advocating that objects also have agency, cause translations and actively take part in sociotechnical networks. The other one, demonstrated with controversy mapping, is that all the elements in a network will always display more complexity than the whole, totalizing view: it is in the details and local connections that complexity lies. But when we consider maps as media objects, for their collective and performative aspect, we realize there is yet another important shift, that is the fundamental sensibility to the fact that subjective exchanges, like in information flows, encodings and decodings, do describe the positions, agreements and disagreements of people, and their shifting place in collectivity and how their engagement in debates takes place. And the visibility to the processes by which this entanglement is constructed might develop new forms of engagement.

Publics are gathered around issues, and this cartographic process of *issuifying* discourses inside controversies changes public debate, if only by giving them tools with which to relate, in the sense of terms, positions and other elements: maps of meaning. And it seems to be a reasonable procedure: contrary to many traditions in the social sciences, controversy mapping does the best efforts to avoid closing controversies into hegemonic positions that would say the truth about the affair, and seeks to create shared tools for debate where actors and relations are defined from traces of their interactions, and not by vertical categories. It is not meant for social intervention, even though it will have a looping effect. For example, we have no doubt that the Emaps atlas on climate change adaptation will popularize the issue of adaptation and the objects mapped to a larger public, changing the controversy. But it will do so mostly for the sake of a shared and inclusive public debate. This signals to the importance of considering the evolution of the controversy, and of considering the effects of mapping to the controversy.

The discussion goes deeper: Marres (2005), while analyzing the possibilities for an object-guided democracy, discusses the idea that democracy happens, not when everybody has unlimited access to uncontested truths, but when institutions fail and information is controversial. The author reminds us of some ideas proposed by American philosopher Dewey (DEWEY & ROGERS (ed), 2012), when she advances that democracy works when debate is unavoidable, and when there are issues around which publics can gather. Like Marres summarizes: “no issue, not public” (2005, p.7). That is probably why it was relevant for Venturini, Ricci *et alii* (2013) to title their article on strategies for engaging the publics of controversies with the maps “Designing controversies and their publics”: this term

"designing", is not aiming exactly at graphic design, but to the idea that publics and controversies are being outlined and mutually constructed through mapping.

5.2.2. Methods and field description

For our fieldwork we were able to follow up on the research activities of the médialab of Sciences Po Paris[5], especially the work done in Emaps[6], one of their main projects. The médialab was founded in 2009 by Bruno Latour, with the purpose of being an innovation center for the advancing of the digital humanities. A staff of more than twenty researchers, developers, designers and other professionals collaborate in research projects to advance social analyses, tools and methods that draw from the technologies of communication and information.

For this research, we had access to staff meetings and documents and able observe and participate where possible in two of the 2014 sprints: the Amsterdam (March 2014) and the Oxford sprint (April 2014), and to follow up project meetings at the médialab in Paris. We also had the opportunity to interview eight scholars, developers and designers who took part in the project and talked about its goals, challenges and the role of visualization in defining and presenting issues, while working for a better understanding of controversies and fueling public debate and citizenship.

Our main concerns were: [1] understanding the method of controversy mapping, its roots, developments and present challenges; [2] understanding how the participants aim at equipping citizenship and public debate, given that this is one of the objectives of the method of controversy mapping; [3] Directly observe their practices to identify their main challenges related to visualization and the general development of an inquiry in controversy mapping.

The persons we managed to interview were: Tommaso Venturini, chargé de recherche of the médialab; Anders Munk, professor at the University of Aalborg, visiting professor at the médialab Sciences Po; Axel Meunier, project manager at Emaps; Mathieu Jacomy, researcher and developer at the médialab Sciences Po, one of the creators of the open source software Gephi; Donato Ricci, design lead at médialab Sciences Po; professor Richard Rogers, director

[5] http://www.medialab.sciences-po.fr

[6] http://www.emapsproject.com/blog

of DMI (Digital Methods Initiative), University of Amsterdam; Peter Gerry, researcher at Young Foundation; and Michele Mauri, designer at Density Design, Politecnico di Milano.

With these general goals and profiles in mind, we developed a basic script for the interviews, that could be adapted, depending on the profile of the interviewee.

The general script for the interviews

Profile

a) Please tell me a bit about your experience, how it led you into your current work, and how is your work nowadays.

b) What are your current concerns in terms of research?

c) What is your take on the participation of each research center that collaborates in Emaps, in terms of the specific knowledge and particular points of view they bring?

Emaps

d) What is your take on the format of Sprints that is used in Emaps?

e) How would you compare the sprints that went on so far?

f) It seems that in the last sprint in Milano you would be mostly focused on making some sort of atlas from the work in previous sprints. Do you already have some ideas on how to structure this atlas, how should it be?

g) How do you relate your center's work with controversy mapping? Do you feel there are differences between the methods you use and the point of view of different scholars that take part in Emaps?

Mapping methods

h) How do you see the difference between data visualization and information visualization? What is your take on network maps (nodes and lines) as visualizations?

i) One of the goals of Emaps project is creating tools for citizenship. How do you think the maps themselves can or should collaborate towards these goals?

j) At the sprints, I noticed some efforts for the use of visualizations for presentation, but also for data processing and analysis. How would you compare the visualizations produced in each stage, inside the context of Emaps and in broader terms?

k) There are many different forms of interaction with maps. From what I could observe, there is a tendency to use interactive resources such as data filtering and zoom in the processing and analysis stages, and static maps for presentation. How do you see this distribution of resources according to these main purposes?

l) Static visualizations, as in print media, are still more portable and stable, in the sense of being encapsulated in a single file, and for that they may present some advantages over interactive visualizations. How do you see the comparison in terms of the distribution or circulation of the maps?

m) How do you see this tension between static and interactive maps in terms of differences in user experience?

n) When inquiries like the ones in Emaps are developed, some believe that if there is a clear story to be told the results will be stronger, so this story starts being developed since data gathering and it presents the final findings. How do you think visualization and maps participate in the structuring of facts towards the organization of stories and their display?

o) Have you ever analysed or displayed real time data? What is your take on the possibilities of those for research and the challenges it poses to map making?

p) And what about data journalism and visualization? Did you ever collaborate with news outlets? What is your take on the use of visualization in data-driven news?

The field of controversy mapping

The ongoing projects at the médialab aim at very varied controversies, in varied societal contexts. At the time of this research, there was, for example, a project called The Fabrique de la Loi[7] (The Law Factory), which produced various analyses about the legislative process in french parliament, organizing and mapping data about proposals and amendments. This, of

[7] For general informations about the project, refer to http://www.cee.sciences-po.fr/fr/recherche/larticulation-entre-participation-democratie-et-gouvernement/la-fabrique-de-la-loi.html. Results were published at http://www.lafabriquedelaloi.fr

course, generated academic papers, but also a tool for public access that let common citizens explore and develop a better understanding of the disputes and controversies present in lawmaking. It is not a manual that explains how french legislation works and describes the main events on the subject. It is actually an effort for giving many actors (including other members of society who are affected by those disputes), a map on the transformations of laws as they are being stabilized, which turns into a visualization of the debates and negotiations of the legislative sphere of french government. Instead of discussing the finished and stabilized artifact of the law, one is able to have an access to the many interactions and translations involved in its creation, an access into the complexity of laws.

Emaps was an european project that aimed at advancing the method of controversy mapping, in the sense of bringing other actors into the mapping process. The chosen themes were climate change adaptation and aging in Europe, very relevant controversies that affect the scientific community, but also society in general. Apart from the médialab, took part in the project the Digital Methods Initiative (DMI) of the University of Amsterdam (UvA), the Density Design Lab of the Politecnico di Milano, the Young Foundation, a british non-profit entity for social and applied innovation, Barcelona Media, a spanish non-profit entity for innovation in digital media, and the Institute of Spatial Planning of the University of Dortmund. The collaboration of these centers also entailed an interesting exchange in methods and tools, mainly DMI's research on the Digital Methods, the information design methods developed in Density and controversy mapping as developed in the médialab.

The project started in November 2011, and lasted three years, ending October 2014. For its last phase, it focused on the theme of adaptation to climate change. Work was organized in four sprints, which were work weeks where scholars, developers, designers and issue experts joined forces for a few days of intense work and gathered, treated and visualised data while developing and improving research questions, issues and analyses. Of course there was much work to be done in between sprints, like trying to anticipate necessities like specific datasets, evaluating results achieved so far and general planning. The result has recently been published in a website[8], following the spirit of an atlas, with the collection of maps associated in what they started to call issue stories, explanations about the methods used and reference to source data.

[8] http://www.climaps.eu

According to Meunier (personal communication, April 10th, 2014), project manager of Emaps, its main concern is to understand if maps of controversy can be used as inclusive tools for public debate, around which more people could gather and participate. Emaps follows the main objectives of Marcospol, a previous project that was developed at the médialab, but while the later surveyed refined digital tools for controversy mapping, it was during Emaps that the method moved fully into the digital technologies (Venturini, personal communication, October 9th, 2014).

5.2.3. Discussion

Following project Emaps was indeed a rare opportunity because it enabled us to observe some of the main research strands that intersect in the method of controversy mapping, some of the most challenging aspects for its future developments and also some valuable information on the tools used and the approach on visualization. The research strands were quite clear by observing each participant institution and what they brought to the table: the médialab brought a strong STS tradition to deal with scientific controversies and the roots in the ANT principles, as well as, of course, the method of controversy mapping. On the other hand, DMI (Digital Methods Initiative, University of Amsterdam), brought on a strong work inside the digital methods from the point of view of media studies, a combination which helped develop a critical posture towards biases of devices and platforms where the data came from and a careful attention to the narrative aspects of the inquiry and, as Venturini (personal communication, October 9th, 2014) points out, making the complexity of the controversies legible to a wider public. The designers from Density Design (Politecnico di Milano) brought on a solid work in information design, participatory design and, as Rogers (personal communication, May 9th, 2014) highlights, an innovative approach on visual data analysis. Finally, as Munk (personal communication, April 16th, 2014) explains, the Young Foundation brought on the specific experience with collaborative work with stakeholders, from their tradition of research in applied social innovation. This would be crucial for the aim of integrating the actors of the controversy in the mapping process.

This leads us to discussing one of the main challenges of the method, which is the participation of stakeholders or actors of the controversy or, still, engaged publics. We have seen that, for controversy mapping and other close variants in the field of digital social research, the constitution of a public can be a very dynamic thing, because it cannot be

described by demographics, but by local interactions and engagement on controversial issues. Therefore, in the case of Emaps, they searched the actors that were more engaged in the controversy of adaptation to climate change. According to Venturini (personal communication, October 9th, 2014), by looking at the preliminary network graphs that were built to track the debates on the web, they could determine that this was mainly an expert controversy, involving climatologists, public policy agents, lobbyists etc. The non-specialized audience would not discuss climate change adaptation, but related subjects like climate change itself and global warming, if it existed and whether it was caused by human activity. Even the debate around efforts for mitigating was present, but not for adaptation. So it was primarily an expert debate, and the organizers made an effort for having a varied group of them present during the sprints. They were referred to as issue experts.

They were mostly incorporated in the project as references for understanding the origins, structure and possible biases of the data used, and for discussing the partial results in iterative cycles of development. For Munk (personal communication, April 16th, 2014), their participation is a big differential of Emaps project, because they can help deploying the controversy and developing relevant and adequate research questions:

> "So if the theme is vulnerability assessment, then a guy who works directly with vulnerability assessment can tell us, in a dataset, what are the problems with that and the interesting discussions about how it is supplied, how it is used. First we would have that person to present it to us, but afterwards it is also important to have that person at hand, to make sure that you formulate the research well… because (...) we have to make sure we do that in a manner that they can recognize."

But, on the other hand, he recognizes some tensions in this relationship: he believes that, while deploying the controversy in a central concern at the médialab, it may not be so for the issue experts. They may be just looking forward to having a very good visualization on a dataset they already know, for example. So the scholars must always be careful not to lose track of the controversy, always search for ways to reinsert it in the process, as they are iterating the research questions. Venturini (personal communication, October 9th, 2014) also sees some tensions in this relationship, some deeper consequences for the interests of issue experts:

> "on the one hand the cartographer, little by little becomes an expert is his controversy, and on the other hand is constantly dependent on the information he can get from his informants. A little bit like an anthropologist or an ethnographer, who would rely a lot on the information they can get his informant. So, this relationship is complicated, for many reasons, probably the most important being that the actors, precisely because they are acting in the controversy, they have an

> interest in the controversy (...), even when they are willing to help the mapper to do his mapping, they also have their interest, so they are also acting strategically in the controversy and in the mapping. And that is something that cartographers have to take into account. For example, trying to have as many and as diverse experts as possible, in order to be sure that he or she is able to represent is his mapping as many points of view as possible."

We think it is amazing how this reflects some issues in journalistic practice, especially what Cabra, for example, tell us about turning the "owners of the data" into good sources in order to be able to translate the data into relevant reports. Indeed, Meunier (personal communication, April 10th, 2014), envisages the possibility of integrating journalists as a different kind of expert or actor of the controversy, mostly because of the ways in which they approach public debate and they build narratives. Venturini (personal communication, October 9th, 2014) identifies many similarities between controversy mapping and data journalism, and believes that controversy mapping could offer much value for data journalism by bringing them their methods of using digital traces, while the latter could bring to the table their experience in reaching larger publics.

The insertion of issue experts in Emaps and the plans for collaboration with data journalists still do not exactly build an approach that would fit lay people and controversies that are not technoscientific. But, on the other hand, we believe it performs a very positive movement of proposing a more collaborative attitude towards the actors of the controversy or the engaged publics. This changes the criteria for the construction of visualizations and can generate a rich base for future efforts of addressing other kinds of controversies and publics.

About the tools used, they work with varied data methods for extracting and treating data, including mining, web scraping and many others. For that, besides custom scripts, the médialab and the DMI have developed some tools, and both of the centers have sites in which they offer these for free. The main ones would be Navicrawler, that crawls connections of links between sites and builds databases for constellations of websites, and Hyphe, along with a series of converting and parsing tools. For the visualizations, the most robust software would be Gephi, an open source initiative that is bred from the tradition of social network analysis. Traditionally Gephi exports static images, but, as the interest in web-based visualizations grew, some tools for deploying HTML/javascript based networks were incorporated to their catalog. Manylines is their most recent visualizing software, also open source, that aims to address specifically this issue, of easier web sharing, also offering better

possibilities for integration with information systems and incorporating interactive features. To this set, DMI adds textual analysis tools and some other specifics.

Indeed, visualization is an important part of controversy mapping: according to Munk, this is in the very roots of controversy mapping and, at the same time, brings a whole new sense to scientific inquiry, because it allows for the participation of a wider range of specialists and actors, and becomes a space in which the inquiries become visible and are collectively developed. Meunier, for example, believes visualizations should come into the research process as early as possible: they should be specified and streamlined iteratively, together with the research questions. The traditions of digital methods involve the heavy use of visualizations in the exploratory stages, but these are in general conducted inside specialist software like Gephi, that offers many tools for exploring the data landscape, but demand a process of translation in order to distribute the visualizations. Either you have static pictures of what went on inside the software, or there must be a specific work of building more streamlined versions: they might follow the definitions reached during the whole exploration process, but, technically, are mostly made from scratch. So one of the many challenges regarding visualizations in controversy mapping, is to better integrate these two moments, of exploration and presentation, in terms of the process of producing the visualizations.

When considering the circulation and distribution of the finished maps, most scholars we interviewed were somehow cautious with the idea of interactive visualizations, especially with real time data. They understand that static visualizations are more portable, and therefore more stable and less exposed to compatibility issues. Of course, static visualizations can deploy controversies quite well, without the need for data filtering, zooms or pans and it is true that they present clear advantages for keeping their structure and presentation across many devices: one can view a bitmap file in several different screen sizes, project it in a wall, print it, mail it, transfer by FTP, bluetooth, email and so on. On the other hand, a single interactive visualization, even if it may be unstable, can offer many views on the data landscape, privileging an exploratory approach. Munk, for example, sees it as an important path to follow. Ricci (personal communication, June 18th, 2014), on the other hand, worries (quite reasonably) about the experience of the user. He explains that, when too many interactive features are added, we may start looking at a complicated cockpit of commands, that demands much learning from the user, and a shift on the design work towards software development that might hinder the focus on deploying and exploring the controversies.

Indeed, as we are going to discuss later on, the issue of interactivity and of how people may interact with either static or interactive visualizations becomes more relevant as the method of controversy mapping aims at reaching beyond the walls of academia. In the published result of project Emaps in Climaps.eu, some visualizations had interactive features, others were static, but this depended mostly on the character of each dataset and on the specific iterations and transformations that came about in each inquiry.

6. FIFTH CHAPTER: FROM ANALYSIS TO PRESENTATION: REIFICATION OF ISSUES, REENACTION OF INSIGHTS

In this chapter we are looking at how visualizations are used throughout the more exploratory stages of data analysis and how they aid in outlining issues and objects, that are also discursive resources for communication and building up insight. We understand that visualizations in the exploratory part of the process differ from the ones used to present findings, but we believe it is more productive to consider this difference, rather than as two separate and stable levels of visualization, as something that is developed throughout the process of analysis, much as a chain of transformations. We would like to contend that this process advances by developing clearer understandings of the issues being studied, which could be related to the hardening of facts and objects of study, parallel to organizing possible connections between these elements in a largely narrative way. Of course, in social research as much as in news reporting, the outlining of issues to be approached is an important part of the whole process, and it can be done through images or not. Nevertheless, our goal here is to discuss how visualizations of data and information, once they are used, can carry out these tasks from analysis to presentation, and to identify some ways in which they make visible the processes of building knowledge and thus can be used for public debate. We will discuss the many transformations of visualizations throughout the process of analysis until the presentation of results. Even though our cases will be mostly from Emaps project, we will trace some parallels with some strategies in data journalism.

As we have indicated in the first chapter, the discussion about scientific images, the many devices used for scientific investigation and the translations from the beginning of the inquiry towards standardized and measured objects that ground scientific discoveries is central in STS literature. In an article titled “Les vues de l’esprit” (1985), Latour had already put forward the idea of the *mobiles immuables*, or immutable mobiles, which would be in the center of scientific activity: normalized and stabilized objects that could be recombined and transported to different contexts. Later, in the book Pandora’s hope (1999) he also beautifully describes a series of transformations between the complex (or raw) object of study and progressively compatible and standardized elements. In the same journal as “Les vues de l’esprit”, Lynch (1985) had already presented an article that focuses more specifically on the process of selection and mathematization of visual documents, that is, scientific visualizations. Later on, Latour & Biezunski (2005) will deepen this discussion by emphasizing the social aspect of

developing scientific knowledge: they will talk about the hardening of facts, that is, many methods and processes by which facts are outlined, proven and reified, many times, to an extent in which they become irrefutable. This connects STS with the sociology of knowledge, and to the implications of scientific images as as social objects and even media devices, as we have signaled throughout this work. Even though all these works highlight some very important issues for our discussion, they are mostly concerned with the whole scientific process and the many devices used and produced, or focused on scientific visualizations, but not on specific problems related to data visualization.

The discussions about scientific images that we identified in STS literature in general consider an initial *messy* object to be studied, like the jungle in Pandora's hope (LATOUR, 1999), that is either measured, pictured, sampled, and the processes of normalization and mathematization that sprout from there. As we described in the first chapter, this is especially visible in the examples used by Lynch (1985) and in scientific visualization more specifically. Nevertheless, in a sense, data can also look like a messy jungle, an ever-growing mass of seemingly scattered givens that we feed with our interactions and with our work, that needs to be parsed, selected and transformed to be visualized or fed into information systems in order to become useful or at least meaningful. In the first chapter, we described a continuum between data visualization and information visualization: from visual displays that are more *raw* and closer to data, to displays that incorporate the results of analysis and aggregated objects. This labor of developing analysis that is paired with refining visual displays bears much similarity with the process of mathematization of scientific visualization described by Lynch. Both will work towards visually defining objects and pulling information from noise. Nevertheless, contrary to the visual documents discussed by Lynch, data itself does not have specific spatiality: even geographic data might be visualized in many different ways, depending on the choices made for analysis. With data visualization, in a sense the image is already mathematized from the beginning, when it is mapped in a geometrical plane even as we recognize there will still be an effort of refining information from it. Also, as we intend to demonstrate, along the transformations in data and information visualization, the spatial quality of the initial visualization will probably not be kept throughout the transformations: new structures and methods may be tried in order to advance the needs of the inquiry. So the chain of transformations involved in data analysis entails many choices on how to select and display its dimensions and variables in space. And we can add still one more difference: while in scientific visualization a pattern or quality should be outlined and observed across a series

of images, with data visualization each graphic also condenses its own series, even if it is part of a series of graphs, and is taken by itself as a context where patterns can be outlined. We understand that the specific theme of the transformations of data visualization along inquiry merits specific attention.

In the fourth chapter, we discussed some aspects of the controversy mapping method, and how it was important that the results of the inquiries be published as a set of many different materials: from network graph versions, to views of details, to other kinds of graphs, source data, texts, lists of actors, trees of disagreement and so on. According to Venturini (personal communication, October 9th, 2014), there is indeed a close relationship between controversy mapping and social network analysis, because both share a deep concern for the interactions and relations between actors and their varying densities. So even though there is no obligation to use them, network graphs tend to be the first and more traditional mapping structures to be used in most projects. Other centers of digital social research, like the DMI, from the University of Amsterdam, for example, will emphasize other graphic structures, like sankey graphs or even those usually associated with statistics, like bar and pie graphs. Their approach and treatment of the data, however, will be very different from traditional statistics, and will emphasize categories that emerge in discourse and other quali-quantitative aspects.

Meunier, project manager of Emaps at the médialab, believes that network maps, even though they offer much material for analysis, can be a bit overwhelming and even whimsical, so their complexity should always be countered with other kinds of visualization (personal communication, April 10th, 2014). Ricci, design lead at médialab, also shares this general conception and adds that the process between analysis to presentation involves an effort of "disentangling the graph from the network" (personal communication, June 18th, 2014). He explains that, since the different relations and their density play a central role on the method of controversy mapping, they usually start by making networks, that are complex objects, closer to the data. The network map may be closer to the core of the data, but many versions may be developed along the iterations in analysis stages, and even other kinds of graphs must be tried out in order to communicate with the different actors that are taking part in the process, and in the different contexts, and develop the inquiry. Ricci does highlight that visualizations are also necessary as communication tools for researchers during work: also because of that it is often necessary to translate what has been established into other kinds of visualization or either simpler or partial versions of the network, but translating is itself part of the inquiry. This effort is also part of a process of separating what is noise and what is

information, depending on the situation and the objects at hand, but it also frames the issues studied and how they can be described through the data available, by making them visible. As the work progresses, new objects like categories and clusters are outlined and it is possible to elaborate other ways of treating the data and visualizing it, according to questions and hypotheses that are each time more precise and that match more defined objects. Each version produced incorporates and presents more and more the results of interpretation and analysis. In fact, the aim of controversy mapping seems to be to take on a very bottom-up approach to social analysis through visualization (see, for example, VENTURINI *et alii*, 2013, p.6-10), so the process of outlining controversial issues by assembling and identifying clusters or sets has been quite evident as we observed their work. By performing this progressive work of refining data and visualization, they get closer to information visualization, and get to visualize more aggregated objects in other methods of visualization, in other graphs. These new graphs may be networks or not. So with controversy mapping new objects are produced whose spatiality does not necessarily map back to the original network, while the data itself can be filtered and converted into other structures.

An example of this process is the network titled CO2 Landscape from ISI-WoS: we have a summary of the process of development in figure 6.1 and a streamlined, yet still complex, version of the network for the years of 1960 to 1969, that highlights the clusters of nodes, in figure 6.2. The authors searched for articles containing the keywords “carbon dioxide” or “CO2” in the Web of Science database, and from them they extracted the cited references, to build an initial network map of reference co-occurrence. As we can see in the illustration, this first network is so crowded that it becomes illegible and does not help visual analysis. Then a second version is created by using a force vector algorithm that makes clusters evident. Then, another layer of data is produced from extracting metadata from cited references: the network of references is added with nodes that give new information about the references. Along this path, many elements were defined and put in relation between one another: articles, references, keywords and categories of keywords, presented in the different colors. Also, we can realize how transformations in visualization are intertwined with transformations in data.

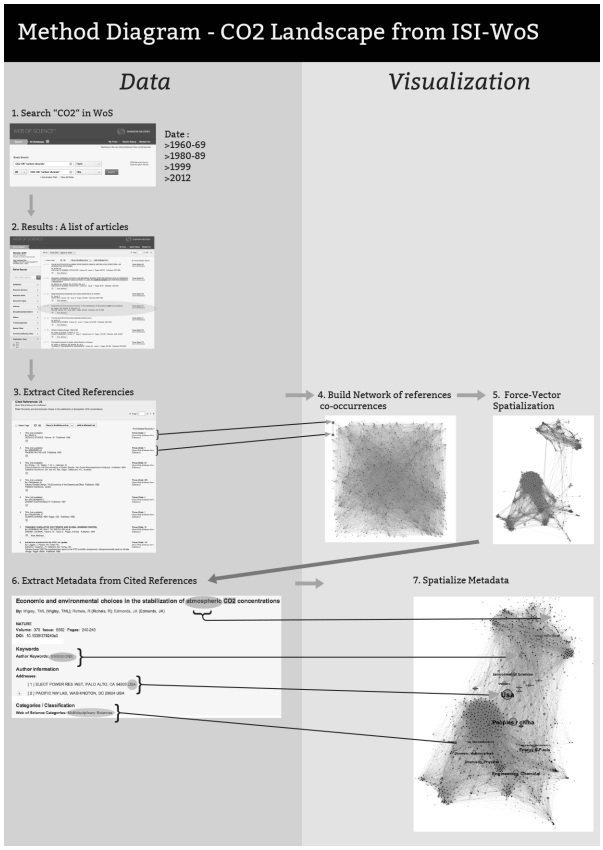


Figure 6.1. CO2 Landscape from ISI-WoS, Method Diagram. By Kari De Pryck & Tommaso Venturini, forthcoming. *See Annex I, figure A1.8, for color plate.*

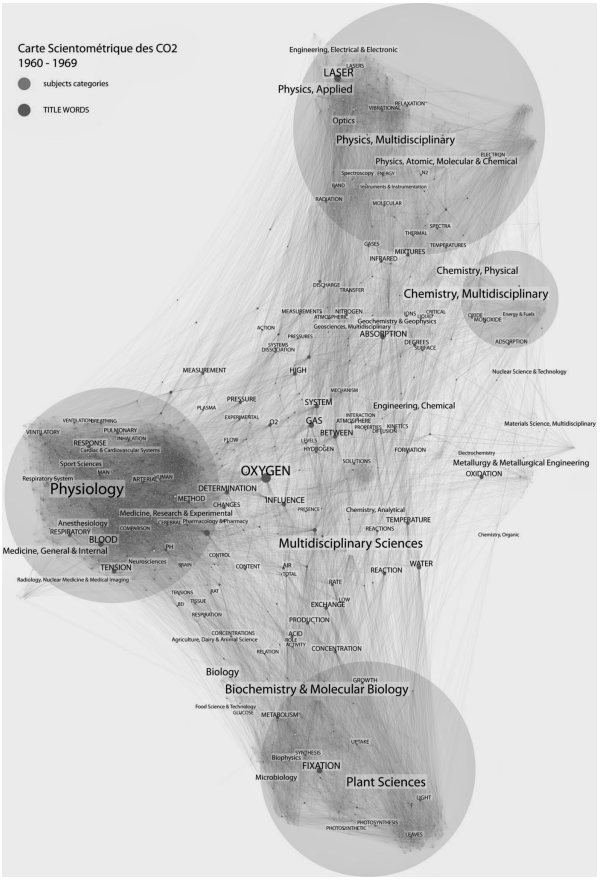


Figure 6.2. Carte Scientometrique des CO2, by Kari De Pryck & Tommaso Venturini, forthcoming. *See Annex I, figure A1.9, for color plate.*

To better examine in the realm of visualization this chain of transformations from initial analysis to presentation, we would like to call it a process of *reification of issues*. We are using the term issues (and not scientific objects or even facts), because we want to highlight the possibilities visualization might offer for public debate. This movement, of reification of issues towards analysis and presentation is complemented by another movement that we would like to call the *reenaction of insights*. This last is related to the concerns about presentation, that will of course entail varying levels of didacticism and rhetorics (OFFENHUBER, 2010). Our point here will be to discuss how insights that happen in the analysis stages collaborate to outline scientific and discursive objects and have to be reenacted (or demonstrated) in presentation, so presentation might, from this path, also benefit from an exploratory perspective.

We should point to the complementarity between reifying issues and reenacting insights. This last is related to the concerns about communication and presentation, that will of course entail

varying levels of didacticism and rhetorics. The many transformations on visualizations are a fundamental part of a narrative of the analysis. This process may also appear in the final results in order to feed discussion, and not just as, well, the final clearly displayed and delimited results. Offenhuber points to a rhetorical maneuver used when researchers demonstrate results of visual data analysis:

> "...after puzzling the audience with a complex visualization, the presenter selects, seemingly arbitrarily, a single data point and connects it to a story, an anecdote that unlocks the principle of the whole representation. I suspect this single data point is seldom as arbitrary as it might seem, in fact the whole visualization might be designed to highlight this single point – a rhetorical device allowing the audience to reproduce the discovery of meaning in the data." (OFFENHUBER, 2010, p.370)

This makes the case of the author for what he calls a "visual anecdote", a strategy that contextualizes the visual exploration in a wider social environment, pointing to its applications and also aiding persuasion by mirroring the audience's direct experiences. We would like to add that this rhetorical strategy also points to *reenacting* the transformations carried out during analysis: the transformations between visualizations, carried out either by making series of static visualizations or by operating interactive ones, are reaccessed, selected, beautified and repurposed if one aims to demonstrate the story behind the results. This is related to what Latour (1999) calls the chain of reference of scientific artifacts, where at each step of transformation it is important to keep the link to the previous step in case results need to be reproduced, but involves a staged presentation, a reenaction of the chain as a story. Of course, in scientific circles demonstration of results is a traditional concern, but with data analysis and especially in controversy mapping, as showing different points of view and aspects of the data is paramount, the issue stories could work towards this recollection of chains of visualization.

In this sense we could say that experimentations with ways in which to present complex data landscapes and highlight different aspects in order to reenact insight are crucial to data journalism, because they seem to endure much more pressure to be clear to a wider and more varied public, while, on the other hand, must face more limitations in terms of schedule with the short deadlines of newsrooms. As scientific methods like controversy mapping try to reach beyond the walls of the academia, they are facing similar challenges. Nevertheless, to locate the reenaction of insights only on the final results would be too narrow: newsrooms, as well as research labs like the médialab, gather many different people, with different backgrounds, so the visualizations during analysis are also instruments for debate, debate

itself being part of analysis. As Ricci points out, there must be a concern for communication between different actors taking part in analysis, and visualizations are the main tools. We understand that, in bigger or smaller amounts, even during analysis, reenaction of insights is rehearsed, and is part of the debate that constitutes analysis and reifies issues.

We should reassess the discussion presented in a previous chapter, about the use of connotation and denotation as analytical categories to deal with the moment of encoding and decoding of messages according to Hall: curiously, the same demonstration that connects data points with familiar narratives in order to reenact insight, is equivalent to the maneuver of connotation in the decoding moment. The reenaction of insights is what bridges specific encodings with more polysemic and controversial aspects of the display that may feed debate.

On the other hand, this reenaction that problematizes the representation itself seems to run counter to the challenge of communicating complex affairs, to which simplification in the representations is usually in order. This is, as we saw, a very important concern for media outlets, and becomes even more important as the publications aim at larger publics. All these transformations inside the exploratory stages towards presentation might work into the frame of simplification, but we would be narrowing the discussion if we mistook this to be mainly a didactic effort: simplification happens a by-product of defining objects.

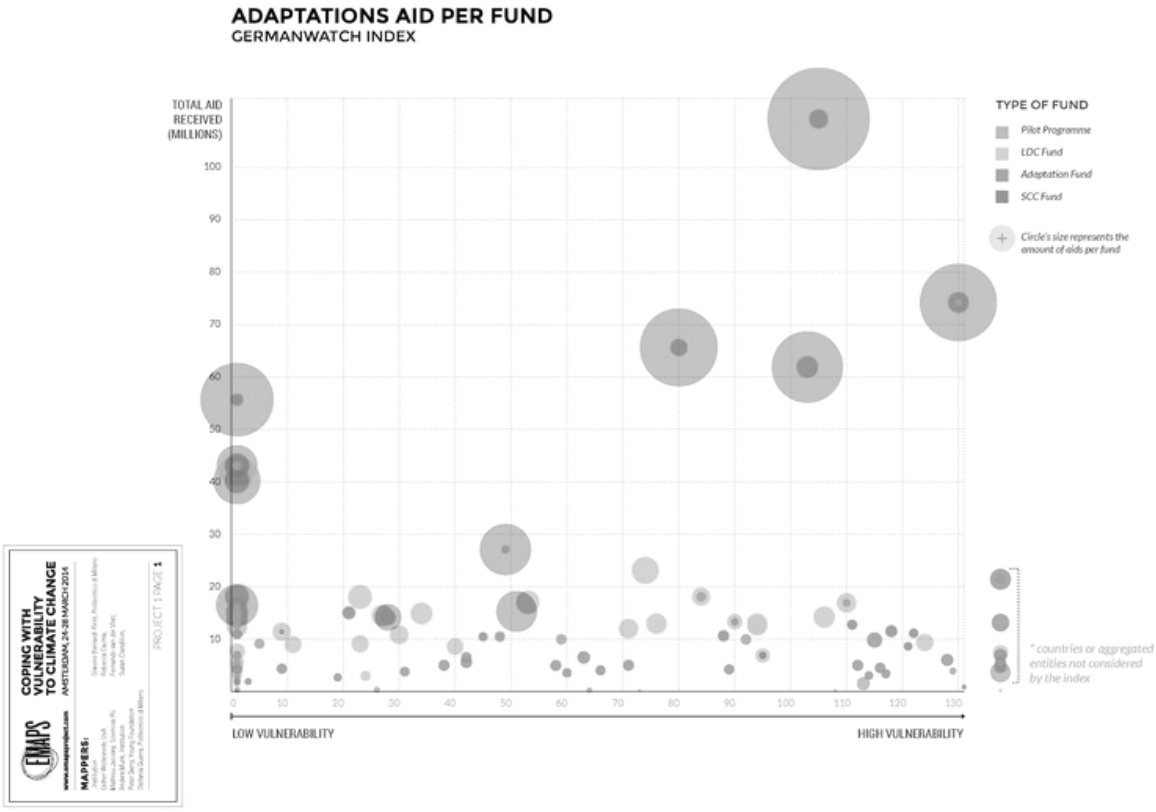


Figure 6.3. Adaptation Aid per Fund - Germanwatch Index. One of the visualizations in the series that displayed the allocations of each fund across countries, ordered by vulnerability to climate change, according to the Germanwatch Index. Circles along the y axis refer to countries that were not indexed. The allocations are represented by their value in dollars. Source: Emaps archives.
See Annex I, figure A1.10, for color plate.

Therefore, we must clear out a classic pitfall when talking about visualizations and the challenge of instrumenting public debate: the idea that simplification is a safe path for making visualizations more broadly accessible. This line of thought and practice can be found in various circles and might be more evident in media, especially in news outlets, and in certain efforts for the popularization of science. We would like to contend that, even though simplification most of the times does occur in the process between inquiries and presentation of results, it should not be considered a necessity or an objective in itself. We think simplification should be seen as one of the resources available, and sometimes almost a byproduct, for the two purposes of *reifying issues* and *reenacting insights* as the work moves toward presentation. This is because the process that goes from analysis to presentation entails of course the definition and hardening of facts and objects of study, as well as the progressive outlining of narratives and rhetorical strategies.

In the case of the presidential elections in Brazil that we displayed in a previous chapter, we cannot credit the problem to short deadlines: as the election day is a planned event of national interest and relevance, the visualizations were projected in advance, and in the main news sites they were quite refined technologically. Some of them even displayed real-time data, as the data from each electronic ballot box was uploaded by the end of the day. We believe that one of the reasons for the problem in the representation of the results was the general concern with simplification in communication, associated with the use of dominant and widely accepted visual formats, but both without a deeper consideration to the socio-political context.

During the Emaps sprints (work weeks of the research project) we were able to follow in detail the transformations in visualizations during data exploration. We identified a chain of transformations starting in the Amsterdam sprint, in March 2014, which we will discuss in the following pages. The sprint gathered scholars, developers and designers from the participant centers in the University of Amsterdam for five days of intensive work. Participants were organized in five workgroups, each of them working with different datasets and sources to explore research questions around the general theme of climate change adaptation. In Group 4, titled Uses and Users of Vulnerability Indexes, there were two projects: the first, aimed at studying and comparing the uses of indexes of climate change vulnerability in the adaptation spaces. In the second project, the team seeked to explore the extent to/ways in which flows of adaptation funds are related to vulnerability assessments. The criteria for adaptation funding had already been a subject of investigation in the previous sprint in Paris, and the group would continue the work.

For the second project, some main questions were defined, and the visualizations we intend to discuss were meant to address two of them:

- Are those [countries] considered most vulnerable also the ones who receive the most adaptation funds?
- Do those [countries] considered vulnerable by some indices receive more adaptation funds than those considered vulnerable by other indices?

The included indexes were: the DARA Climate Vulnerability Monitor, Germanwatch's Climate Risk Index and Maplecroft's Climate Change Vulnerability Index. The UN Human Development Index was also used, once it was identified in the debates that were analysed a perceived link between vulnerability to climate change and lower human development

conditions, that led to precarious living situations, and thus a higher exposure to climate hazards. A basic dataset on the allocation of funds by country was available for each of the major international funds: the Adaptation Fund (AF), with a total of $189,720,000 allocated in the timeframe analyzed; the Least Developed Countries Fund (LDCF), with $504,379,998; the Special Climate Change Fund (SCCF) with $179,280,000; and the Pilot Programme for Climate Resilience (PPCR), with $413,200,017.

The adaptation indexes were turned into rankings and two kinds of bubble graphs were tried out: one was a static image generated from vectorial graphics. It fitted the whole plane in a single glance by superposing data points of the allocations of the four funds (figure 6.3). We could have a clear idea of which countries were getting more or less funds and compare their vulnerability according to each ranking, one at a time. The other graph was based in HTML and Javascript, and was structured as a long list of countries organized by vulnerability, with circles proportional to the amount of funding received, relative to the total budget of each fund (figure 6.4).

At the end of the sprint, as all the groups presented their results, the HTML graph, however problematic in terms of information design it was, got a lot of praise from DMI's director Richard Rogers. Later on, we had the chance to interview the professor and asked him why did he approve it so much. After all, it had many design problems, we did not have a general view of the data points (scroll was needed) and it had less dimensions, failing to show the actual amounts of funds being allocated. He just answered that the graph answered very well the question at hand. As we were puzzled by this reply, we made one more desperate attempt: was it because it is interactive? He just said that the graph was useful because showed very clearly that the indexes were not a relevant criteria for the choices of the funds regarding their allocations across countries (personal communication, May 9th, 2014).

Indeed, looking closely, one of the differences of this graph in relation to the others that were produced at the same time, from the same basic data set and with the same basic visual structure, was that it displayed the amounts in terms of the percentages of the total budget each fund allocated. So we did not have a visual representation of which countries were getting more money, but of the chosen priorities of each fund, to be compared with the ranking positions according to each index. This allowed for different considerations, more geared towards discussing criteria of funders and the relevance of the indexes, and not the financial result, and was, indeed, more connected to the question at hand.

In the final results that were published in the site Climaps.eu, the two approaches are present: in figures 6.5 and 6.6 we see different views of an interactive version of the initial static graph in figure 6.3, that shows the fund's allocations in single circles for each country, avoiding superposition. Now the colors of the bubbles are only used to reinforce the growing vulnerability assigned to the countries towards the right. In figure 6.7 we see a much clearer and condensed, streamlined version of the initial HTML graph in figure 6.4. Each visualization presents, of course, different aspects of the data, and the final versions were not exactly simpler than the initial ones.

In the Oxford sprint, that took place at the Centre for Environmental Studies in the University of Oxford, in Abril 2014, participants were gathered in four groups, where two of them used variations of the HTML bubble graph in order to display square matrixes. For Group 3, that seeked to profile adaptation practices, it was turned into a large graph for initial exploration and comparison of data points that described adaptations projects along many topics (we see a partial reproduction of this graph in figure 6.9).

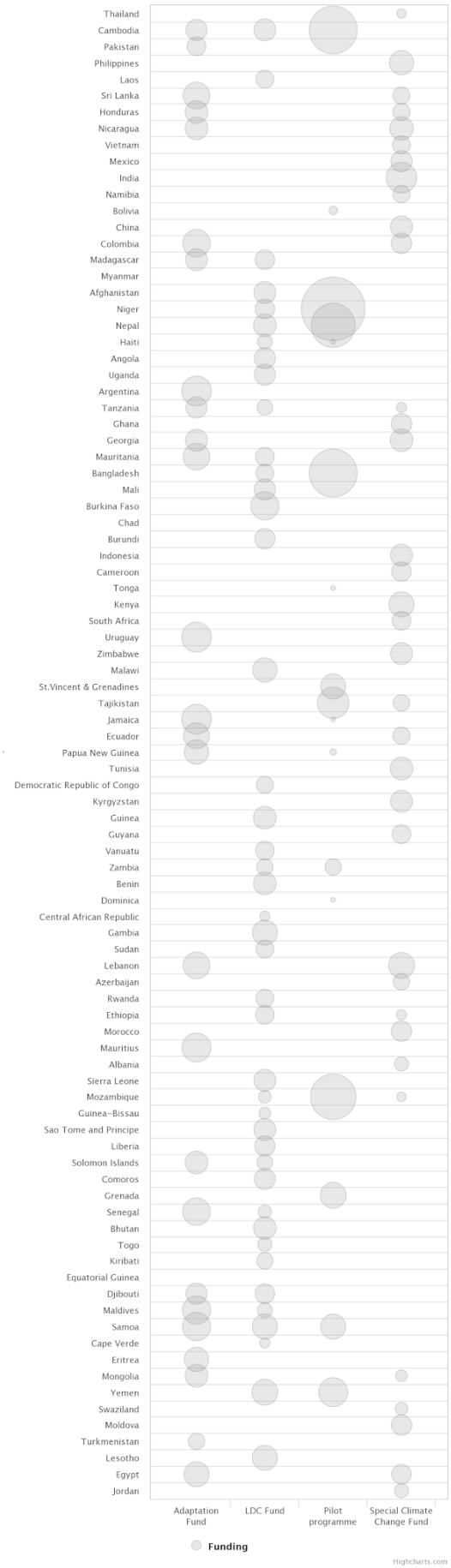


Figure 6.4. A view of the HTML visualization that displayed proportional allocations of each fund for each country, ordered according to the Germanwatch Index.

Group 1 was also interested in profiling projects, but for the goal of understanding which kinds of projects and hazards were more related to the insurance industry. So the same structure was used, but with different data, combining sets from two different sources: the databases from UNDP and ci:Grasp. Both of them are available in websites of public access. The figure 6.8 shows an example of a project profile in the ci:grasp site: the site interface organizes the many fields that describe all the projects in a certain way, which gives a contextualized view of each project but does not allow comparisons en masse, between the more than three hundred projects on the catalog. As the the people in Emaps did not have access to the source database, a custom script was built to scrape the data from all the pages, and the resulting dataset was repurposed into the bubble grid. So the new structure offered a different view on the dataset, a different form of building information from it, while also helping to combine two different datasets, and enriching the landscape for profiling the projects. Of course, graphs that aid in an initial appreciation of large amounts of records, with many data points, will be the starting point to many transformations in data and into new visualizations. This enormous bubble graph was, quite humorously, named “The big grid” by Bounegru and Rastall, two of the Oxford sprint participants.

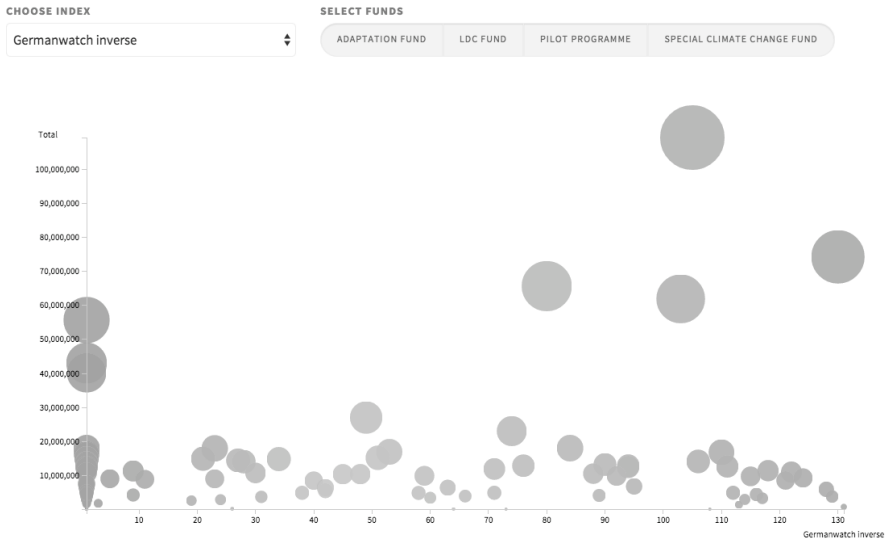


Figure 6.5. Multilateral Adaptation Funding And Vulnerability Indexes. The interactive version, where one can choose an index in which to order the x axis (Germanwatch in this case), and to select the funds to be included in the mapping, thus changing the size of the circles proportionally. Source: http://climaps.eu/#!/map/multilateral-adaptation-funding-and-vulnerability-indexes

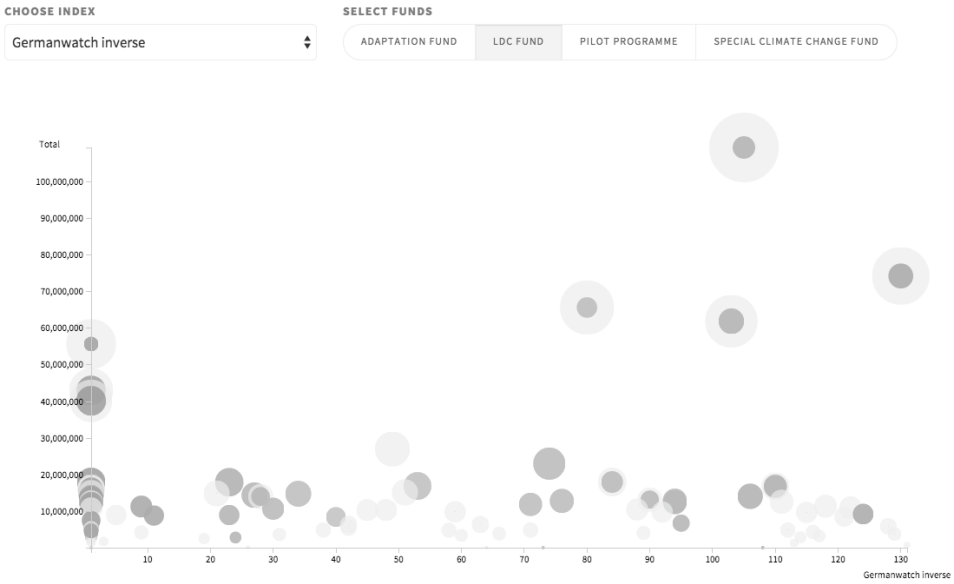


Figure 6.6. Multilateral Adaptation Funding And Vulnerability Indexes. Another view from the previous visualization, displaying only the LDC Fund in the colored areas, with the gray circles indicating the total funds allocated to each country. Source: http://climaps.eu/#!/map/multilateral-adaptation-funding-and-vulnerability-indexes. *See Annex I, figure A1.11, for color plate.*

As an example, according to Meunier (personal communication, April 10th, 2014), afterwards the grid was used to compare the cases of India and Bangladesh, in order to enrich the work of Oxford sprint Group 4, that had the main goal of tracking adaptation finance in Bangladesh. This comparison generated the drafted graphs in figures 6.10, 6.11 and 6.12, and was represented in a more complex interactive visualization in the final results in Climaps.eu (figure 6.13). The drafted graphs (figures 6.10 to 6.12) still used the language of bubbles to represent the different proportions at each variable, but the final interactive visualization at Climaps.eu replaced them with bars, that had the advantage of showing the proportion of each value to the maximum value displayed, aiding comparison. Nevertheless, the general table-like structure continued, was inherited from the square matrix of the bubbles (figure 6.9): now it is a square matrix of bars.

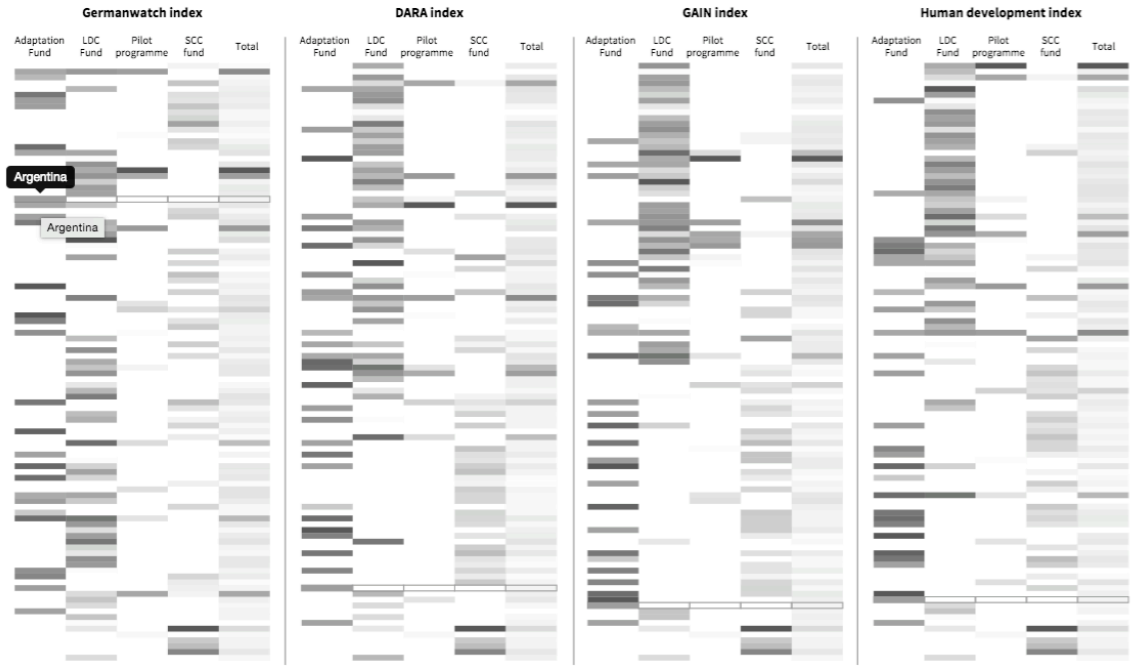


Figure 6.7. Multilateral Adaptation Funding And Vulnerability Indices - Matrix. In this version, the bars for each country are ordered vertically according to each index, and the color density is proportional to the amount allocated from each fund. An extra column was added to represent the total amount allocated for each country, hinting at the ones that were more heavily financed in absolute numbers. In this print, we are highlighting Argentina, that is assigned as quite vulnerable by the German Index, but not at all according to the others. Source: http://climaps.eu/#!/map/multilateral-adaptation-funding-and-vulnerability-indices-matrix. *See Annex I, figure A1.12, for color plate.*

We identify three main movements of transformation in this case: first, the display of the first bubble graph was refined into a more streamlined and clear presentation; second, the structure of the big grid was a bit polished in order to function as a simple tool for other research questions; third, the analysis of the data in it drove new data treatments and the production of other, more advanced and precise visualizations (see a schematic summary in figure 6.14). We believe that it becomes quite clear that, across work groups and sprints, from analysis to presentation in the final atlas, simplification was not the driving force, neither for reifying objects, nor for reenacting insights.

While examining these movements, we initially come across two interpretations for the idea of visual simplification. First, we could think of simplification as a process in the chain of transformations in scientific images and devices, as described in STS literature. It can be understood as a progressive schematization that condensates many relevant cues and information, and, thus, visually simplifies an initial messy object. Second, we could think of simplification as a concern for communicating, as part of the effort of making visualizations

more broadly accessible. This second interpretation for the subject of simplification is present in the classic literature of information design, like, for example, when Tufte talks about data ink and improving the data ink ratio of quantitative graphics: he advances we should clear all the visual information that does not display data and would be, therefore, decorative.

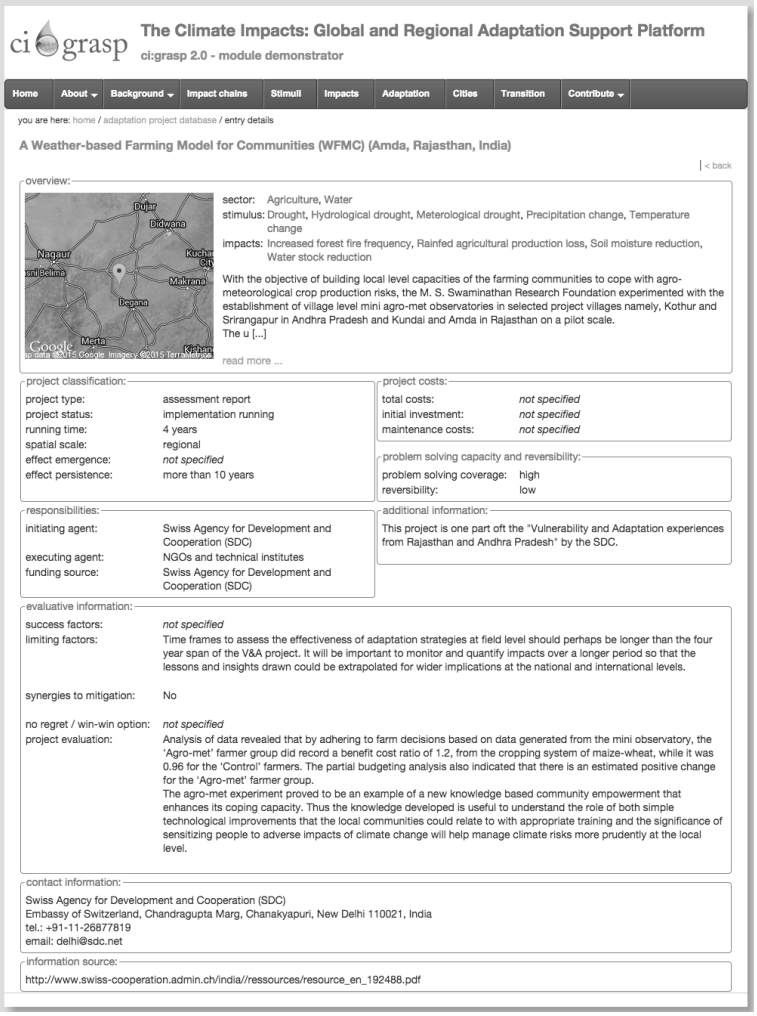


Figure 6.8. A project profile in ci:grasp website. Source: http://www.pik-potsdam.de/~wrobel/ci_2/adaptation_projects/325.html

While both facets of simplification are present in the transformations we examined, we would like to advance that there is yet a third aspect to these movements which we will call *shallowness*. In the first movement, the html bubble graph, despite its many problems in terms of information design, did display very clearly the (non)relation between the priorities of funders for distributing the available funds across countries and the vulnerability indexes. We believe this is related to the fact that it had a very flat presentation, where the bubbles were distributed almost as if in a table, with no superpositions, no depth in the presentation; and, at the same time, a normalized dataset, where differences in the sizes of total budgets were cleared out of sight by treating them as percentages. Since the research question at hand was geared towards the priorities and choices of each fund, and to comparing those to the

priorities suggested by vulnerability indexes, displaying amounts in dollars would introduce unnecessary depth to the visualization, considering the funds had very varied amounts of money. In this sense, the initial bubble graph is shallower than the better visually designed visualization in figure 6.3, because, on the one hand, it is more limited and more focused, and, on the other hand, it is more regular visually. We consider the final streamlined version in figure 6.7 to be even more shallow, because it does not depend on navigation to display all the data for all indexes, organizing them side by side and aiding comparison. At each step, the visualizations did not become simpler, and not necessarily displayed less information.

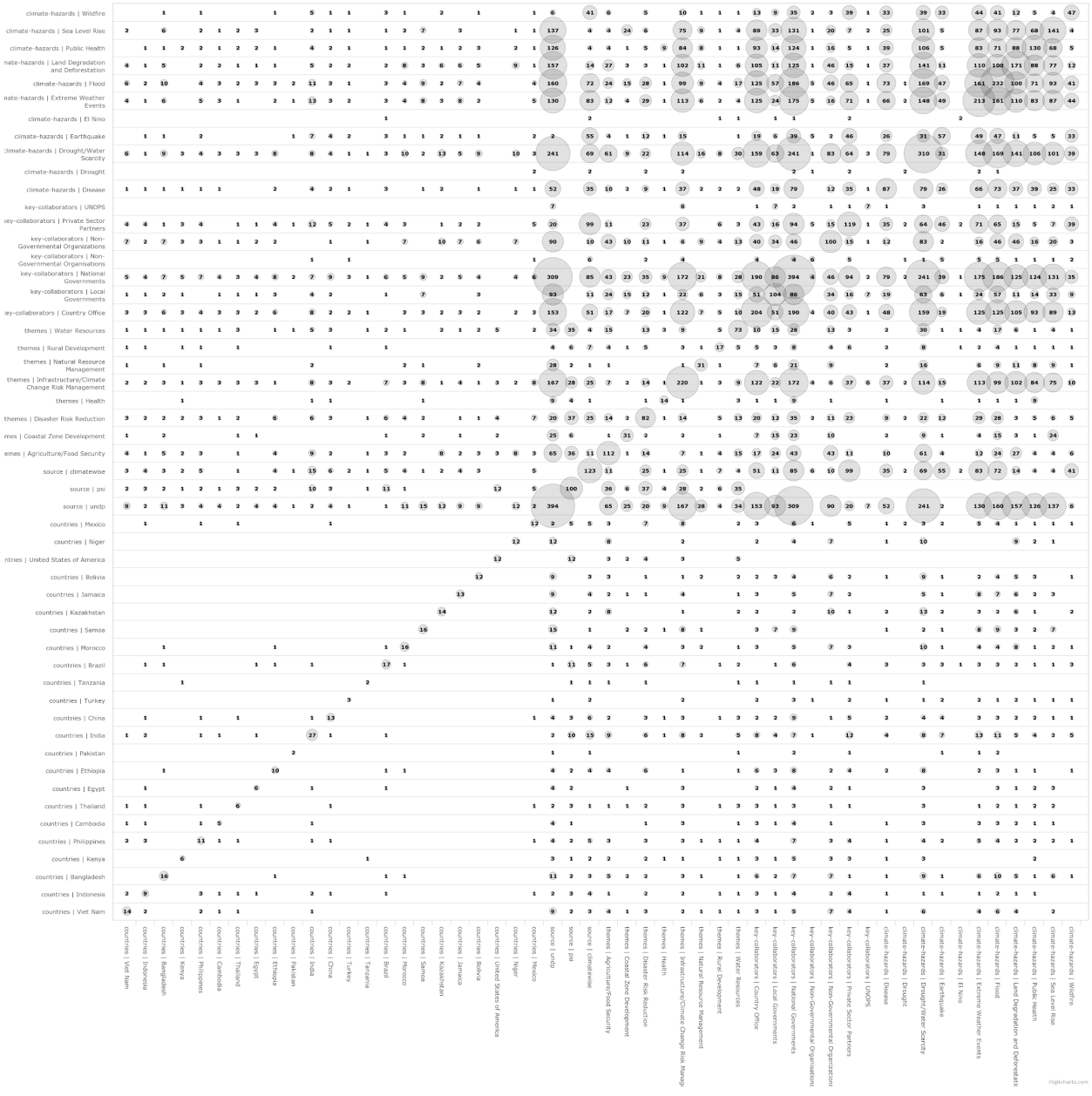
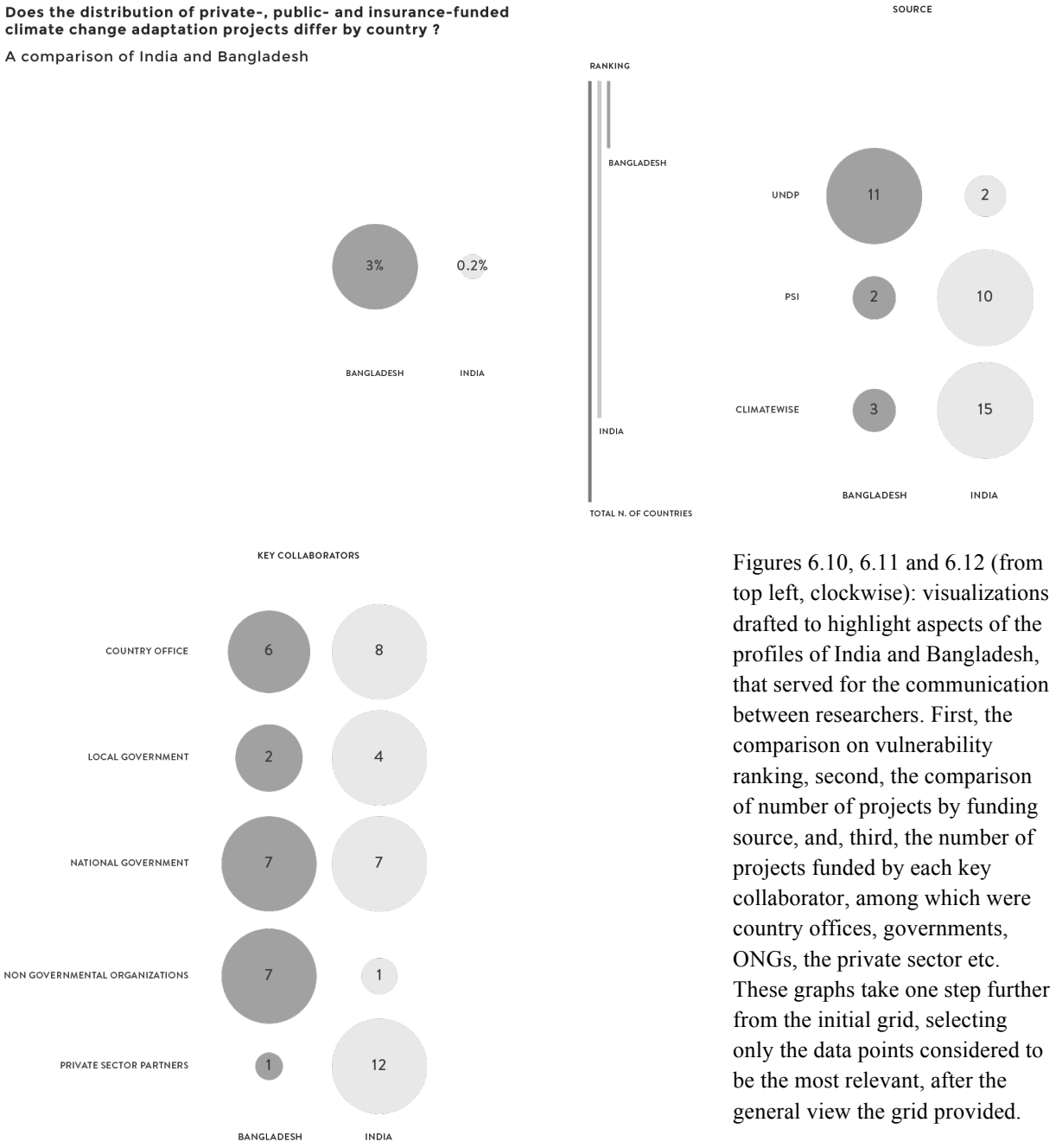


Figure 6.9. *The big grid*, a square matrix to integrate datasets from different profiling sources and to help identifying patterns between many projects.



Figures 6.10, 6.11 and 6.12 (from top left, clockwise): visualizations drafted to highlight aspects of the profiles of India and Bangladesh, that served for the communication between researchers. First, the comparison on vulnerability ranking, second, the comparison of number of projects by funding source, and, third, the number of projects funded by each key collaborator, among which were country offices, governments, ONGs, the private sector etc. These graphs take one step further from the initial grid, selecting only the data points considered to be the most relevant, after the general view the grid provided.

In the second movement, the bubble graph was turned into a square matrix for profiling adaptation projects: this time, on the contrary, despite some new information design problems, some extra elements were added to the structure to ease exploration so it could be a general tool for integrating different datasets and to an initial exploration of data. In the third movement, the structure was used to reconcile different datasets so that researchers could have a more complete landscape, and afterwards specific data points and details that were

identified in the grid were displayed in more limited visualizations, that advanced the labor of outlining issues and enriching their description.

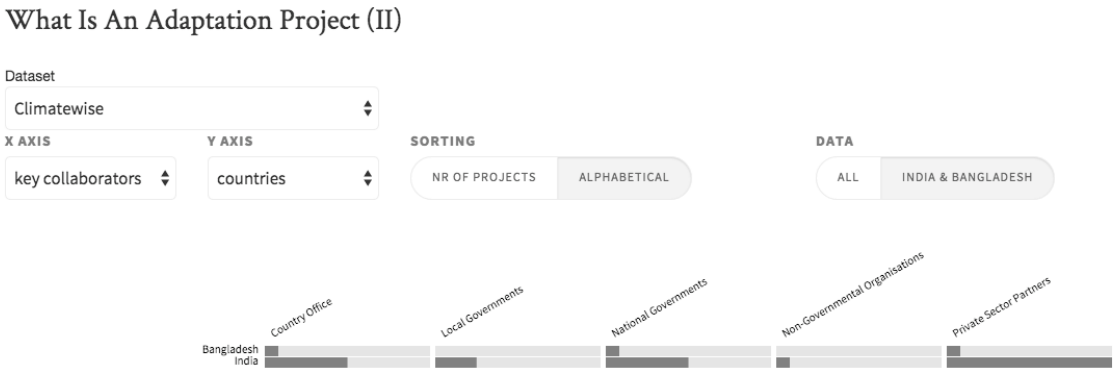


Figure 6.13. One of the visualizations that were finally published on the site Climaps.eu, on the subject of adaptation project profiling and using the same datasets: a tool where the user can explore different combinations and either verify the findings exposed in the text or search for other meaningful comparisons or patterns. There is a highlight for the India and Bangladesh comparison, but there is also the possibility of taking a broader view, choosing to see the data on all the projects from the countries available.

Network graphs, as explained by Venturini (personal communication, October 9th, 2014), display relations between nodes and their proximity and density, and are probably the more direct visual representations of the vision of ANT. Nevertheless, for the questions at hand in the Oxford sprint, that involved different levels of adaptation project profiling, and for the data available on these projects, it was interesting to use the format of the square matrix, in order to be able to compare their characteristics. So we should note that the big grid (figure 6.9), that was produced from the initial bubble graph, became something like a general tool that was used for an initial appreciation of data, just like the tradition of the network graphs. It is closer to the data, displaying a raw interpretation of it, and the transformations and conversions that followed came to *disentangle* (Ricci, personal communication, June 18th, 2014) other graphs that are more interpreted, as the result of further analysis and iteration. The main criteria for choosing between using the big grid or a network map for an initial appreciation will be mostly its matching to the research question at hand and to the structure of the datasets available. Of course, using both resources will offer different views on the data, will present different descriptions to advance the research questions.

By looking at the diagram of diagrams that we composed in figure 6.14, we also realize that, as in every scientific inquiry, the work on Emaps is deeply interconnected, and this is shown in the transformations between maps. Maps that at the start were aimed at one set of

questions, became tools to address others; those that started as series (that are traditionally called small multiples) were transformed into interactive versions, either displaying each multiple according to a set of configurations or flattening the variety of views in one single surface; others were drastically simplified with a narrow focus and then complexified again. We see also variations in grid use, as between figure 6.3 and 6.4, and conversions between grid-like structures to table-like, to list-like, and many combinations in the final interactive ones. Many other visualizations were produced and transformed during the sprints and, as one can verify in the Climaps site, many were published. Due to the highly collaborative profile of Emaps, there are probably many more influences and transformations to be traced.

In and all, neither simplification nor didacticism seemed to be defining issues for reaching the final presentations. Effective communication, in this project, was reached by combining the reification of issues with reenaction of insights: the final maps communicate by demonstrating. Surprisingly enough, scientific verification is found to be an interesting communicational asset for introducing complex structures, in such a way as to generate many understandings through exploration and discussion.

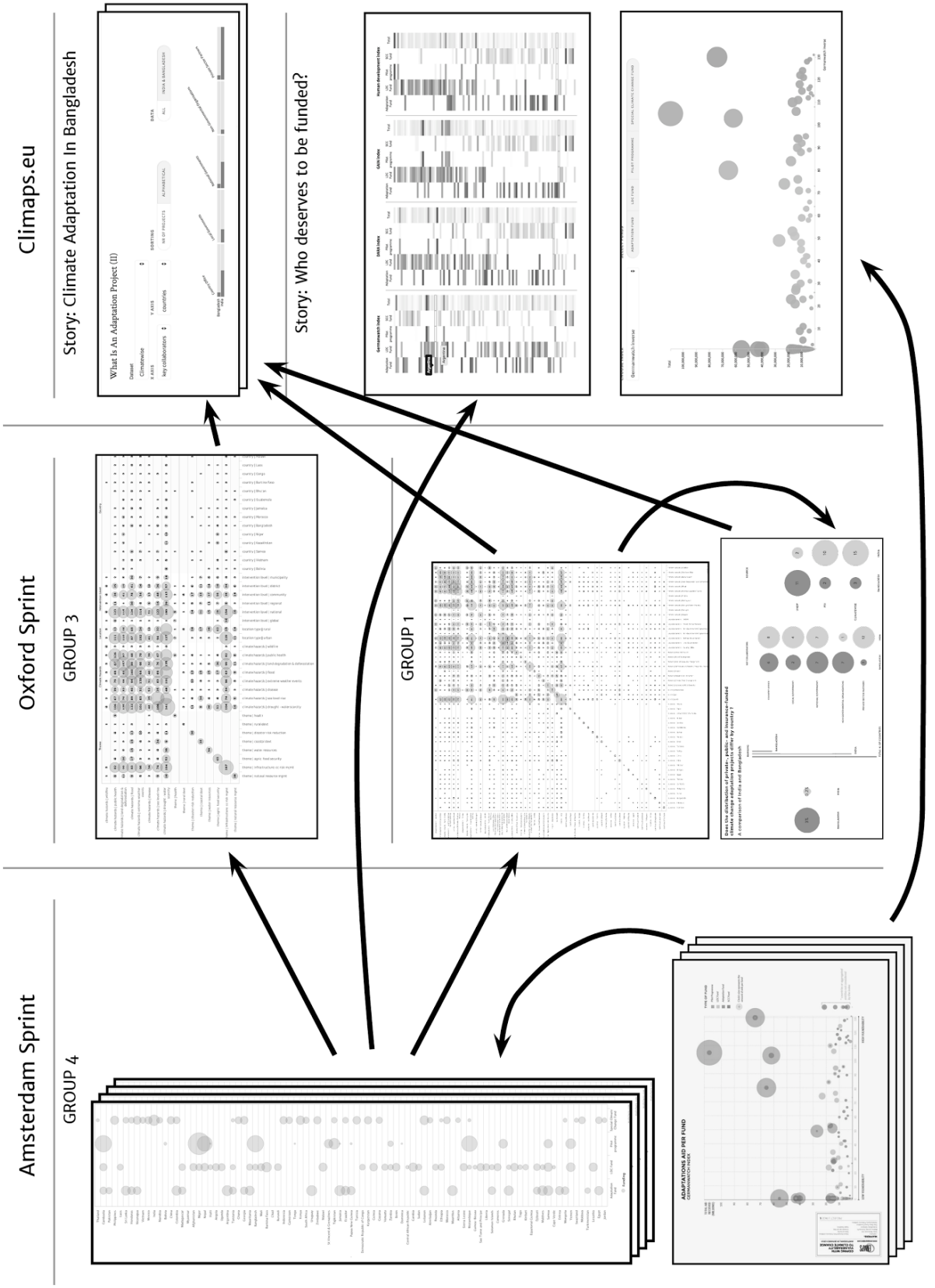


Figure 6.14. A schematic representation of some of the transformations observed in Emaps. See Annex A, figure A1.13, for color plate.

7. SIXTH CHAPTER: SORTING IT OUT AND NAVIGATING: THE CAMERA AND THE ALBUM

"I think the best way for us to look at these maps would be like an animation", says Peter Gerry, researcher in Applied Social Innovation at Young Foundation, as he shuffles between beautifully crafted bubble graphs, displayed on the big screen. It is the final presentation, the closing of an exhausting and very inspiring week of work in Amsterdam, and Gerry presents some of the results of his workgroup. The participants of the work week (called sprint) include scholars and professionals from different research centers in Europe, data analysts and information designers, who worked in 5 different teams to develop maps on diverse aspects of climate change adaptation.

Each graph offers a selection of data points of the same data set, but their joint narrative was only visible or approachable in the format created for presentation. This is an evidence of how analysis and presentation are not separated stages, but should be taken as dimensions, many times coincident, of inquiry, but it also highlights the importance of how the access to visualizations, and the narratives involved, are also determinant for the inquiry. In this chapter we would like to follow the thread of the complementary concepts of the reification of issues and the reenaction of insights to propose procedures by which visualizations are actually organized and accessed in the process. For that, we propose yet another pair of concepts, which will be related to two complementary approaches: the approach of the camera, in which exploration of data landscapes is carried out through composing different views from visualizations; and the approach of the album, in which several different maps are aggregated to form a narrative. We could say that the camera approach will be more evident during analysis and the album approach at the presentation. But these forms are neither mutually exclusive nor restricted to such moments, as we believe the discussion instrumented by these same concepts will help to show.

We will use two main examples: first, the visualization Cidade da Copa (City of the World Cup), produced at the infographics division at GloboEsporte.com, by Carlos Lemos and Fabio Pena, under the coordination of Mario Leite[9]; and, second, the atlas of Emaps, that, as we already described, presented the final results of the project in the website Climaps[10]. The first

[9] http://app.globoesporte.globo.com/copa-do-mundo/cidade-da-copa

[10] http://www.climaps.eu

example will mainly demonstrate the characteristics we are assigning to the approach of the camera, while the second will exemplify the approach of the album.

First of all, we should note that both strategies aim at developing visual narratives as a procedure to display different aspects, objects and issues of a theme. From our point of view, both of them are related to the idea of scientific atlas, as defined by Daston & Galison (2010): collections of images that display objects of scientific interest according to relevant scientific criteria. For the authors, scientific atlases, however varied forms they might have assumed historically, have a deep impact in scientific practices because they provide them with working objects that can be collectively shared and used as references, landmarks for evaluating specimens and evidences. They point out that classic scientific atlases served scientific practice by training the eye to "see the essential and overlook the incidental, which objects are typical and which are anomalous" (p.26), so that every practitioner would look at and refer to the same thing, would be "visually coordinated" (p.26). Atlases, therefore, made scientific collaborations possible: "the atlas is always – and fundamentally – an exemplary form of collective empiricism", they "enroll practitioners as well as phenomena".

Daston & Galison build a history of the many conceptions of objectivity and how the corresponding epistemological ethea were defined and sought after by refining scientific images in atlases, and by disciplining the eye. Like we pondered in the first chapter, the authors aim mostly at what we have been calling scientific visualization: from ideal models in the eighteenth century atlases, that seeked to convey the perfection and regularity of nature in hand-made paintings; to specimen that were mechanically reproduced in pictures where every effort was made to painstakingly efface every trace of human subjective interpretation; to annotated images and series that would display patterns found by the trained eye; they define three main epistemologies of the eye. These are: truth to nature, mechanical objectivity and trained judgment, respectively.

The point made by Daston & Galison (2010) is that scientific images in atlases configure a scientific ethos, in the sense of providing means, the space in which to produce scientific knowledge and moving in scientific issues. They are intrinsically a collective undertaking, they are meant for sharing. This relates to what Hall proposes about the different moments of encoding and decoding, and to the operational maps of meaning and effects of meaning attribution: while Hall stresses codes and maps of meaning attribution, Daston & Galison will stress atlases and epistemological ethea. Either way, for the purpose of our work, we must

point to the performative and collective procedures of the atlases we are discussing, and to the possibilities they offer for setting ground for debate that is intrinsically connected to knowledge production. Differently from the previous chapter, we will not be focusing on the transformations between visualizations, but in the composition of spaces corresponding to these ethea, however flexible and plastic they might have become in contemporary debates.

Instead of scientific visualizations, our atlases gather information visualizations, and the outline we are trying to build is much more fluid: we need a theoretical object that may account, not only for the transformations that took place in its composition, which form the base narrative, but also for the possibilities of many decodings that will come back as many reencodings, that is, further transformations. This is necessary in order to fit our collections in the contemporary social landscape, where debate is closely coupled with information and media devices, and maps become media devices themselves. Both controversy mapping and data journalism should cultivate a sensibility and an attention to these appropriations and further transformations, that are after all part of the digital traces to be reencoded for representations of social life. Our atlases are, therefore, collections of visualizations of data and information (or unique complex visualizations) that temporarily stabilize conflictual subjects, in order to compose a shared field in which debate can occur.

Nevertheless, the collective aspect and even the effort of visually coordinating and of training the eye remains relevant. An atlas will aggregate views, tell a story and mostly guide the reader's visual interpretation. It will perform a narrative function that is two-sided: at once, it recovers the transformations that led to its configuration, thus looking at the past, and yet, it proposes new narratives that relate to the reader/user/actor's experiences, aiming at future possible developments. It will develop a shareable set of criteria for accessing visualizations, therefore aiding to create common grounds for debate. All these tasks can be addressed in many different ways, with resort to many different structures.

Both the procedures of the camera and the album borrow references from cinema and photography theories. The camera approach will focus on framing, zoom, pan, many camera movements. As we described in the first chapter, these movements entail cycles of alternated human interaction and transformations in views, mappings and data, that are the subject of interest of HCI (CARD *et alii*, 2013). It is more closely connected to our traditional idea of interactivity. A device for seeing offers some controls with which we explore a landscape, the montage is dynamic. The album, on its turn, assembles many views, sometimes based in

different structures and different datasets, and demands that one discovers a narrative in between the patterns of each visualization. On the first, one performs the movement of developing a narrative to produce a collection, on the second, one accesses and recombines a collection to build a narrative.

Eisenstein (2002), the experimental movie-maker and cinema theorist, advanced, in the beginning of the twentieth century, that cinema had the potential of being an universal language, and that what made it cross the barrier of simply transmitting continuous recorded flows, towards engendering concepts, was cinematographic montage. With montage, the mainly indicial transfer of scenes into film by means of a camera could be cut into more iconic representations and reassembled into sequences of related objects, that would produce discursive meaning, the cinematographic concept. He advances that, the more iconic the representation gets to be, the more the cinematographic language may develop a grammar that engenders this second semantic level. While proposing these ideas, Eisenstein takes resource in japanese ideogrammatic writing, and not on phonetic writing. So the most important thing to accomplish in order to create a specific cinematographic language was overcoming the indicial origins of its images towards a more iconic presentation and, from there, grammatically reassembling a narrative towards symbolic concepts.

Recently, Lévy (1998) analyses this dream of cinema as an universal language, that most pioneers cherished in its origins, and argues that cinema never could reach this ideal for a lack of linguisticality, that is, for not being able to match linguistic structures. For example, he posits that there cannot be a direct equivalent for words in cinema. He explains that in phonetic writing we have the elements of the first articulation, which are endowed with meaning: these are the morphemes, or words that are combined to form syntagmas. And then we have also the elements of the second articulation, which are the phonemes, that possess just enough value to be different from one another, but do not have meaning. In cinema, the second articulation is absent: because of the fundamental indicial character of its images, the tiniest element, even a single film frame, is already endowed with unique meaning. Therefore, there are no cinematographic words, because there is no filmed image that is not already a scene, a duration, a phrase.

It is clear that Lévy has associated the idea of achieving an universal language, and even of having linguisticality, necessarily to the structures of verbal communication and, more fundamentally, to phonetic whiting. We should note that the ways in which images may be

invested in building messages and discourses may be very varied, and may develop sometimes fleeting but nevertheless complex grammars. There are cinematographic forms and resources that move between the fundamental indicial character and a deeply symbolic and, to an extent, syntagmatic elaboration: that is precisely where its specific narrative form is bred. In this sense, Manovich (2001) develops a very useful parallel between cinematographic language and the language of new media, while discussing how the first prepared some resources that are later enlivened by the second.

For the author, and that relates to the general theme of discretization and complementary reassemblage that we have been elaborating on this thesis, modernization involved a disruption, a fragmentation of of the continuity of physical space, which gives way for producing interchangeable and mobile signs over objects that are rooted as original references. Cognitive experience becomes each time closer to a constant rearrangement of reserves of these compatible, interchangeable signs, much like the montage as described by Eisenstein. So one of the main challenges and potencies of cinema and, at the same time, of digital media, is merging the dynamics of database and of narrative into a communicational form, that will be based in interface operations. We would like to advance, continuing the general approach that we have been building throughout this thesis, that the key for describing the discursive aspect of atlases is in this fluctuation between fragmentation and continuity, of the database and the interface, of the collection and the narrative, that the procedures of the album and the camera perform.

The composition of atlases bring up a discussion about interactivity and the uses of interactive features in visualization, that we believe can be quite misleading. As research in the digital humanities and new forms of journalism turn to the web and experiment with digital networked presentation tools, there is the promise that new interactive resources might bridge the practical and instrumental gap between investigation and presentation, giving more people the opportunity to have an exploratory perspective over data and conveying more information with access to detail, while offering more persuasive ways for simulating the insights experienced during the exploratory phases. Indeed, in some cases we recognize a tendency for the use of interactive and more complex and open visualizations on the exploratory phases, and what would be considered as static visualizations for presenting results. We believe this is mainly because for visual data analysis there exists complex and specialized software, that would provide sophisticated tools for different structures and many transformations in the visual exploration of datasets. Meanwhile, for presentation there are different needs in focus,

such as compatibility, stability and easy distribution, which are in general best addressed with static visualizations. From this organization of things comes the idea of bridging presentation and analysis through interactive resources as a strategy for improving the debate based on data visualization. This would lead us to design concerns other than specifically information design, that are those of interface and software design, for example.

Nevertheless, we believe that, on the one hand, interactivity, as it came to be experienced through digital software, does not always provide adequate resources for exploration or collective debate, while visualizations that are considered static might succeed as tools for discussion and collaboration, and help building collective understandings about controversies. On the other hand, it is necessary to understand that, even though digital networked technologies do bring very interesting and even groundbreaking interactive resources, interaction and interactivity themselves come in all shapes and sizes, as in many technologies, techniques, tools and situations of use, many of which are older than digital technologies, and still applied in many communication devices. Mayer (2011), for example, develops a very interesting discussion about how even printed scientific images can engage much more than the trained eye, and how these many possibilities of interaction make them collective objects of inquiry. We should therefore avoid thinking of the distinction between interactive and static visualizations as a clear frontier that is traced with digital technologies and consider a broader scope of interaction. Even though we do believe new interactive resources should definitely be explored, in order to advance on our specific discussion we should rather try to focus on how the composition of atlases may give way to varied experiences of access and not on these specific interactive features.

In fact, our particular version of the idea of an atlas is quite loose: it comes from the general assumption that, in order to feed debate and avoid reifying previous assumptions or conceptions that reproduce structures of social power and dogmatic knowledge, we should make our best efforts to assemble many different approaches and aspects of the same issues. Atlases are combinations of visual assemblages of data (the visualizations), either to be explored, either to be related.

The following examples, besides clarifying our conceptions of atlas, camera and album, will also allow us to discuss a related issues, which are the relation between text and visualizations in atlases. In the scene we described in the beginning of this chapter, Gerry was shuffling through some of the visualizations we talked about in the last chapter: we had four different

vulnerability indexes that were used to rank dozens of countries, and four different funds and the data regarding their allocations between countries. So one of the strategies to verify if there was significant correlation between assigned vulnerability and adaptation funding was to develop a series of bubble graphs, each displaying countries ordered by a specific index and bubbles sized according to the amount of funding received. As there were four different funds, bubbles were separated into color-coded concentric circles, each color corresponding to one of the funds (see figure 6.3). Of course, there were patterns and significative data points to be discovered and discussed in each one of the of the graphs, but the answer to the research question was in between the collected graphs: it was not single a pattern, but a comparison of a series of patterns, an appreciation that had to move across the series. Interactivity in this sense was missed in between the maps, not exactly inside them. That is probably why Gerry felt the need for an animation, or some resource to aid a comparison of patterns.

In the final atlas at Climaps.eu, the graphs were condensed into an interactive visualization: although it had the same general structure for mapping the data, it transferred to the user the decisions of how to pass from seeing the display of one part of the data to another part (see figures 6.5 and 6.6). The commands allowed the user to select which index would order the country list in the x axis, and also which funds' allocations would appear. This solution, a tiny atlas, was, as most of them, in the middle ground between the camera and the album: on the one hand, it demanded the interaction in the form of a sensorimotor loop, of action and reaction in order to examine the set, and, on the other hand, it displayed a limited and predefined collection of maps that should be compared, of which there was no actual exploration, no structural transformation. Despite its interactivity, there is mostly a collection being shifted through, not a modulation of a data landscape from which one pulls interesting views. While we are considering this visualization to be itself a tiny atlas, we should observe that it was published inside the larger atlas of Emaps, so the concept of atlas we are working on also encompasses the possibility of encasing one set inside the other.

The pair camera/album also brings up the issue of the use of text while accessing or presenting visualizations. In this sense, visualizations that privilege the exploratory aspect and the procedure of the camera tend to have no annotations or accompanying comments or analyses. This is mostly the case with many of the more elaborated interactive visualizations some news outlets have been experimenting with. On the other hand, sets that privilege the assembling of different points of view over the same issue through data (be it through several datasets or different visualizations of the same one) tend to use more text, as to contextualize

methods and integrate the many pieces into a coordinated whole. But once again, these are not clearly separated procedures. For example, Gianordoli (personal communication, June 27th, 2014), information designer at Ideo, points out that people who design interactive visualizations should learn with print visualizations how to annotate more, how to develop contextual annotations that would make exploration more worthwhile. As we have observed in the work of GloboEsporte.com, one of the main challenges they have been taking on is to develop integrated ways to annotate or highlight interesting aspects in complex exploratory visualizations. Along these lines, we have the visualization in figure 7.1, that shows the main visualization, with interactive features for exploration, and a few featured stories derived from it. This solution combines the editorial narrative with an exploratory perspective. While discussing the result achieved with this visualization, the team of designers came to the conclusion that it would have profited from offering the user the possibility of taking and sharing snapshots (Lemos, Pena & Leite, personal communication, may 27th to June 1st, 2014). Of course, social media sharing is very positive for improving a webpage's visibility, but apart from that we could point to another aspect, that interests us more: snapshots would crystalize significant moments of the user's exploration, and help connecting the visualization with other narratives in the wider social context, thus increasing its relevance as a whole by feeding many conversations. Here we see aspects from the camera approach combined to the album approach in the collection of highlights, that after all can be used as entry points for a wider exploration.

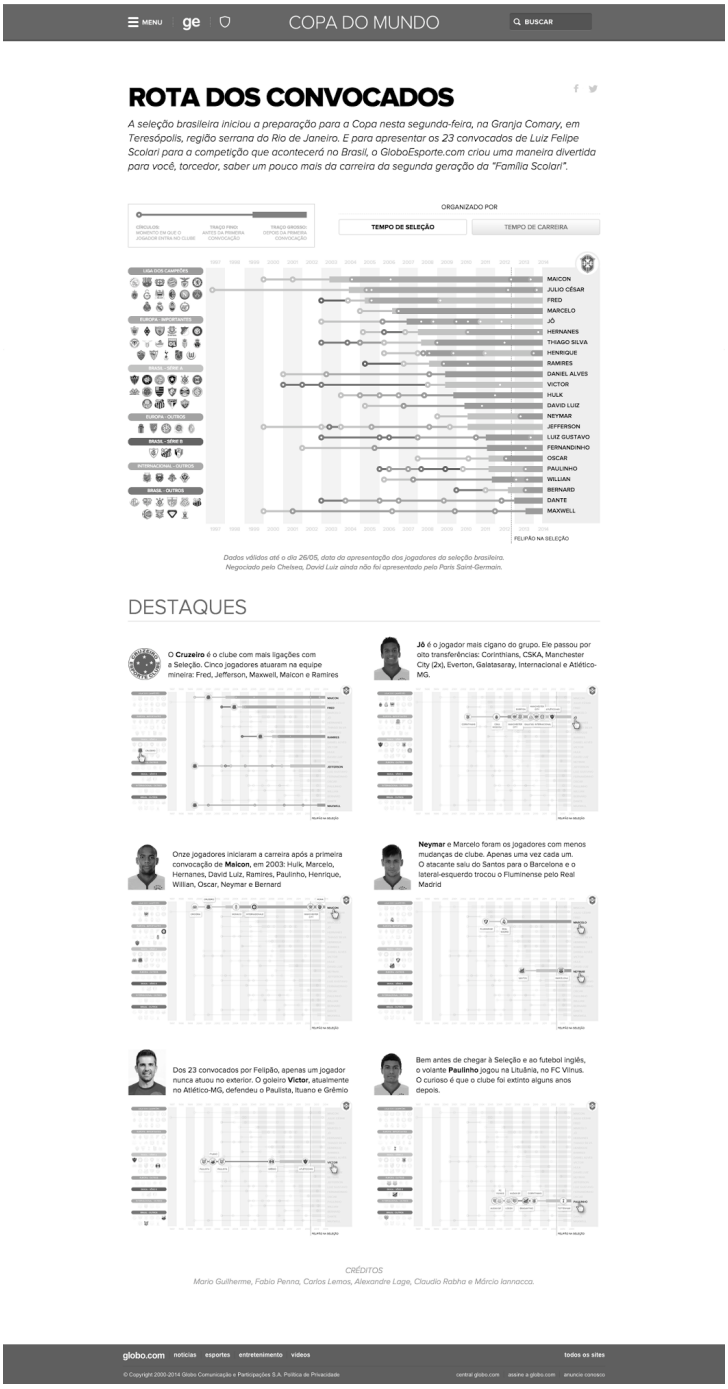


Figure 7.1. Rota dos Convocados, displays parallel timelines, one for each soccer player that was drafted for the Brazilian team at the 2014 World Cup. They indicate and ranked the teams in which each one played, and the duration of contracts. Below, some highlights on specific timelines and annotations developed by the editors. Source: http://globoesporte.globo.com/futebol/copa-do-mundo/rota-dos-convocados.html

But it is probably in print media that we find the strongest examples of the integration between editorial narratives based in text annotations and the visualizations, as described by Gianordoli. In fact, these are quite traditional resources. In figure 7.2, we see a page spread from magazine Superinteressante, a publication that is mainly dedicated to popular science, history and curiosities. We see a combination of different methods of visualization, text analysis and demographical data with annotations. It is a collection of different and integrated

views, using different visual features. This is didactic, but at the same time proposes a discussion.

Figure 7.2. *Quais as últimas palavras dos condenados à morte?* (What were the last words of those who faced the death penalty?) Design: Gabriel Gianordoli, Revista Superinteressante, abril de 2011.

During the 2014 Fifa World Cup, Lemos and Pena, both working at the infographics division of GloboEsporte.com under the supervision of Leite, collaborated to create a quite complex interactive visualization that showed the patterns of goals for each match in the history of the event. They called it City of the World Cup, an interesting but also difficult metaphor that was inspired in the areas of the graph and the height of some vertical bars that referred to the number of goals. In fact, it was not just one visualization, but many that transformed into one another, in a narrated sequence where the user could pause and explore at will (see figures 7.3, 7.4 and 7.5).

We can see here a very interesting combination of guidance and didactics with tools of exploration, and also a mixture of the cinema camera, where the movement is predefined, with some elements of a video game camera, that matches the user's interaction. Instead of integrating paragraphs of text explaining the main highlights chosen in the graph, the camera movements that explored the visualization were organized according to a script and

accompanied by a voice-over narration. The many transitions we see in the animation do not emphasize conclusions or the final work of the journalists or editors, but the structure of the visualization itself. It narrates a demonstration of its geography, or we should say urban landscape, to follow the chosen metaphor. The graphs change and convert into one another before our eyes, showing us many interesting landscapes. In determinate moments, the structure stabilizes and we can make our own explorations. The graphs themselves are quite simple, but the variability of classic structures is explored to the maximum in those passages. We have the tendency to think that exploration may free the reader (or user) to reach their own conclusions, therefore better addressing journalistic objectives, but if a reader has to stop and learn how to operate a visualization full of controls, like a complex cockpit, the issue is not visible. This sort of structure demonstrated the main issues of each related visualization and eased the user into operating it, so the ones who are interested can go on exploring.

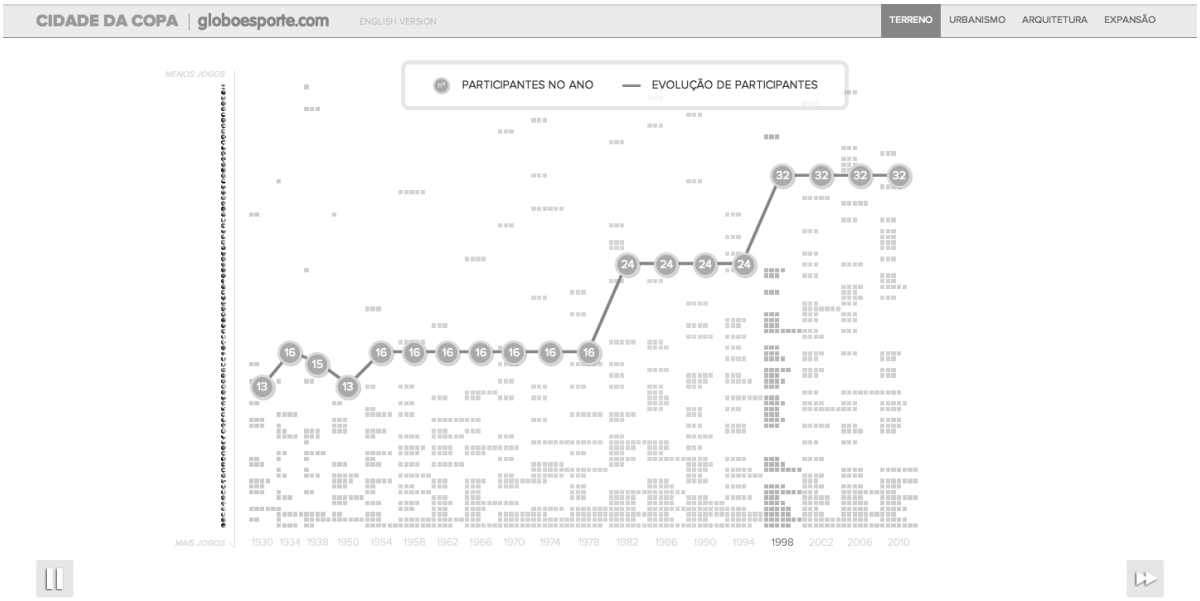


Figure 7.3. Cidade da Copa: highlights along the way clarify the structure of the integrated visualizations. Source: http://app.globoesporte.globo.com/copa-do-mundo/cidade-da-copa/ *See Annex I, figure A1.14, for color plate.*

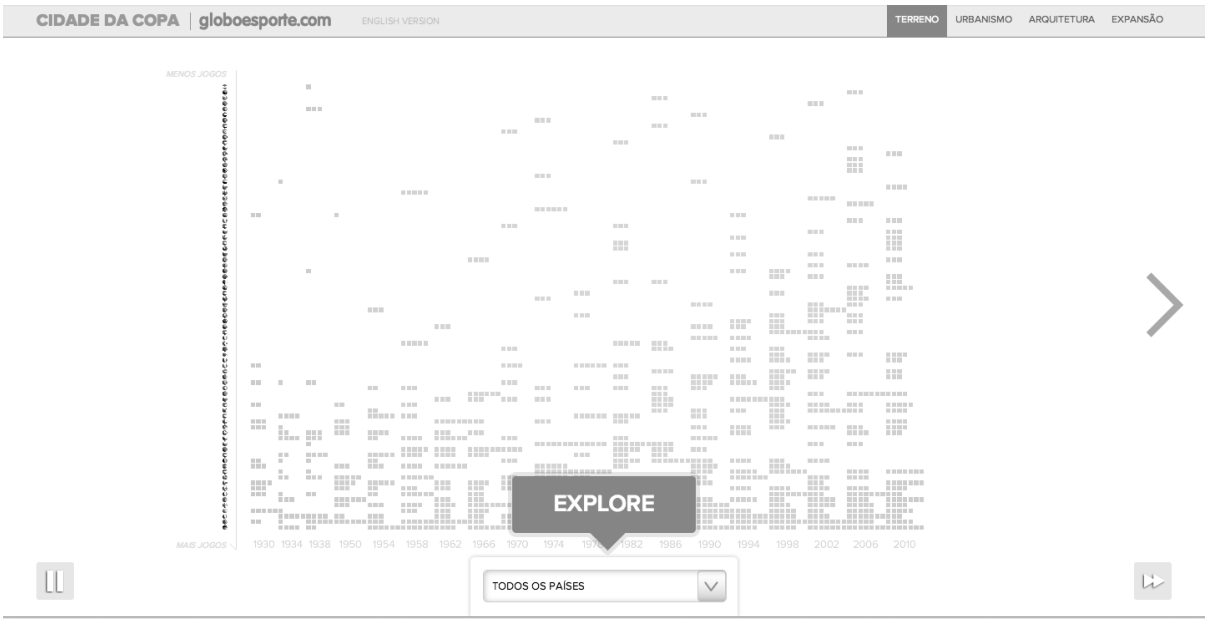


Figure 7.4. Cidade da Copa: In specific moments, the user can explore one of the versions presented or choose to follow with the presentation. All the time there are the buttons for pausing and advancing in the footer. Source: *Idem*. *See Annex I, figure A1.15, for color plate.*

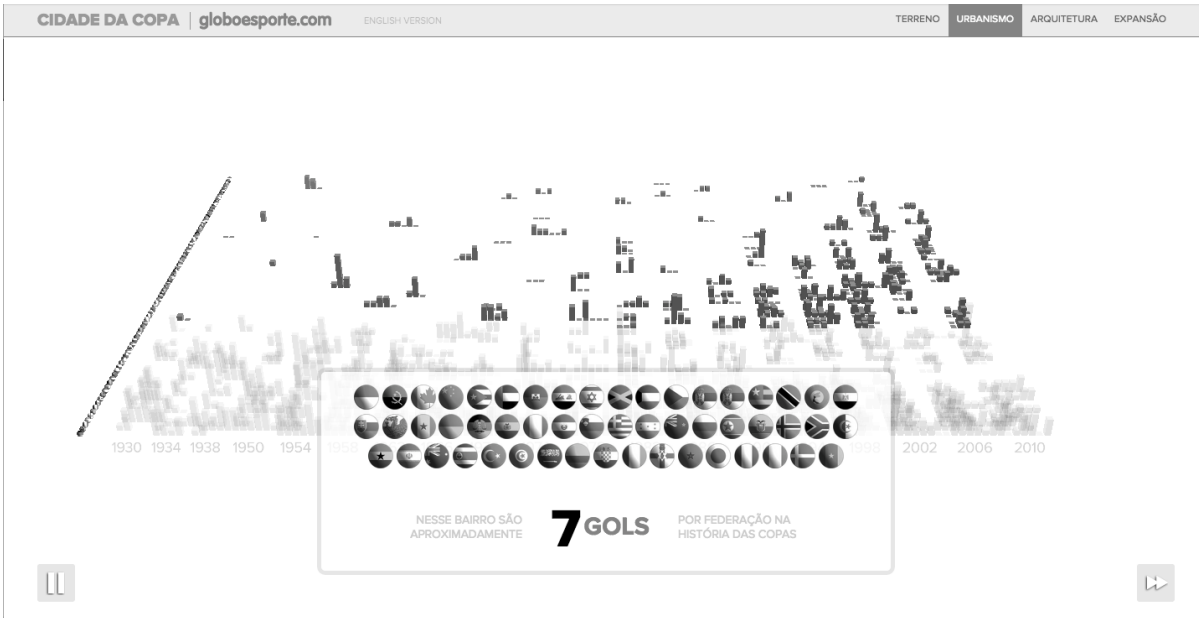


Figure 7.5. Cidade da Copa: This image demonstrates the metaphor of the city: the three-dimensional variation presents bars that are taller according to the goals scored in each match. The visualization is divided into four main structures that accompany the urbanistic metaphor: Terrain, Urbanism, Architecture and Expansion. Source: *Idem*. *See Annex I, figure A1.16, for color plate.*

Pena (personal communication July 27th, 2014), who developed the screenplay, comes from a background in audiovisual media and brings on a different view towards visualization. He remembers that, while working on this project, he and Lemos would often fall into the issue

of deciding whether it was a video or an infographic. Whenever this happened, he would remind the colleague that it had a script, audio narration and animated movement, so it was a video. Nevertheless, the difference from this work and most narrative movies or videos is that the script seems to be bred from the structure of the visualization. Cidade da Copa makes this aspect more evident, but we can see signs of it in all the other examples, as long as there is at least some basic data analysis.

Coming back to our reference in the language of cinema, this reminds us of Peter Greenaway, the movie-maker and visual artist, and what he once said, on the occasion of the centenary of the first commercial cinema exhibition. He declared that even after a hundred years of history cinema was not yet cinema, in general it still had not developed its own language, being too focused on illustrating text and narratives still derived from the model of eighteenth century novels (ARRUDA, 1998). As we consider this radical declaration, we can't help but noticing that with visualization the process might be a little inverted: text and narrative seems to spring from visual cues, not the other way around. It is like the screenplay would come at the end, being constructed along many visual operations and transformations.

Which leads us to our last example, the large atlas that was built for publishing the results of Emaps project. It is a full website, that organizes 33 visualizations, grouped in five issue stories. For each visualization (or map, the term used by the participants), there is a specific page, that explains how it was conceived and developed, gives access to the datasets and sources and, if possible, to the tools used to treat the data and create the visualization. It also states to which issue stories the map is related, for there are some maps that appear in more than one story. Each issue story is also a page, with a longer discussion about the debates addressed, and comparisons and relations between maps (see, for example, figures 7.6 and 7.7). Scholars and designers who took part in the atlas' development point out that they wanted its structure to be as simple as possible, so that the visualizations could take central stage.

From what we observed, this association of many different maps is not a particularity of Emaps, but a recommendation for controversy mapping. Maps of controversy are almost never published alone, that is, only one map published. Even if in specific inquiries one may choose to compose a single complex network, it is going to be necessary to make multiple snapshots of that same network, with different classifications, view of details etc. Normally research conducted in controversy mapping tends to use several different visualizations to

show different aspects of the data and levels of detail and complexity, and combine them with text and other media in order to compose narratives. Therefore we have the controversy website or atlas, that gathers several maps in a digital narrative, while it is also common to develop traditional academic papers.

In the surface, the format that was created for Climaps seems to take reference in traditional scientific papers: objects, hypotheses and research questions, visualizing the material that was collected, discussion, conclusions. Nevertheless, it encompasses some profound issues with narrativity and narrative building that of course are the base to assembling and using atlases in general. First of all, there is the issue of the use of text for narrativity. We notice that in the issue stories and also in the map pages, there is no interference of textual annotations inside the visualizations. The text that appears inside visualizations in only for the labels and, in some cases, in tooltips, little tags that appear according to mouse interactions to offer specific information, like more labels or values. This strong integrity of the space of the visualization does not match examples in journalism, like the figure 7.2, where part of the interest is precisely in how the designer managed to integrate and combine different elements into a harmonious spread. The text in Climaps is, therefore, more continuous and linear, it is not specially spatialized to integrate with the visualization structures. Likewise, the visualizations have quite precise borders and limits.

For Venturini (personal communication, October 9th, 2014), adding a layer of narrativity is essential for engaging the public of a controversy with the maps, and this should be done with the texts, because, to a certain extent, it is only possible to tell stories through text, and it is definitely very hard to tell them through visualization. Also, he believes that a narrative demands well defined objects, and those are outlined in analysis, so telling a structured story is only possible at the end of the research project. Although there is definitely a narrative being unfolded in the analysis stages, along all the exploration and the insights, it is implicit. So for presentation it must be organized and made explicit, in the additional textual layer. In this sense, we can say that the issue stories are the narratives of how the issue was mapped.

The issue of exploration in the final presentation comes up: Venturini believes that offering some possibility of exploration of the results, either with interactive features or not, is important in order to deploy the controversy for the public. Once the story is told in the text, the public can go further and explore the complexity of the controversy. Of course, this perception matches the logic behind the combined solutions in Cidade da Copa, for example,

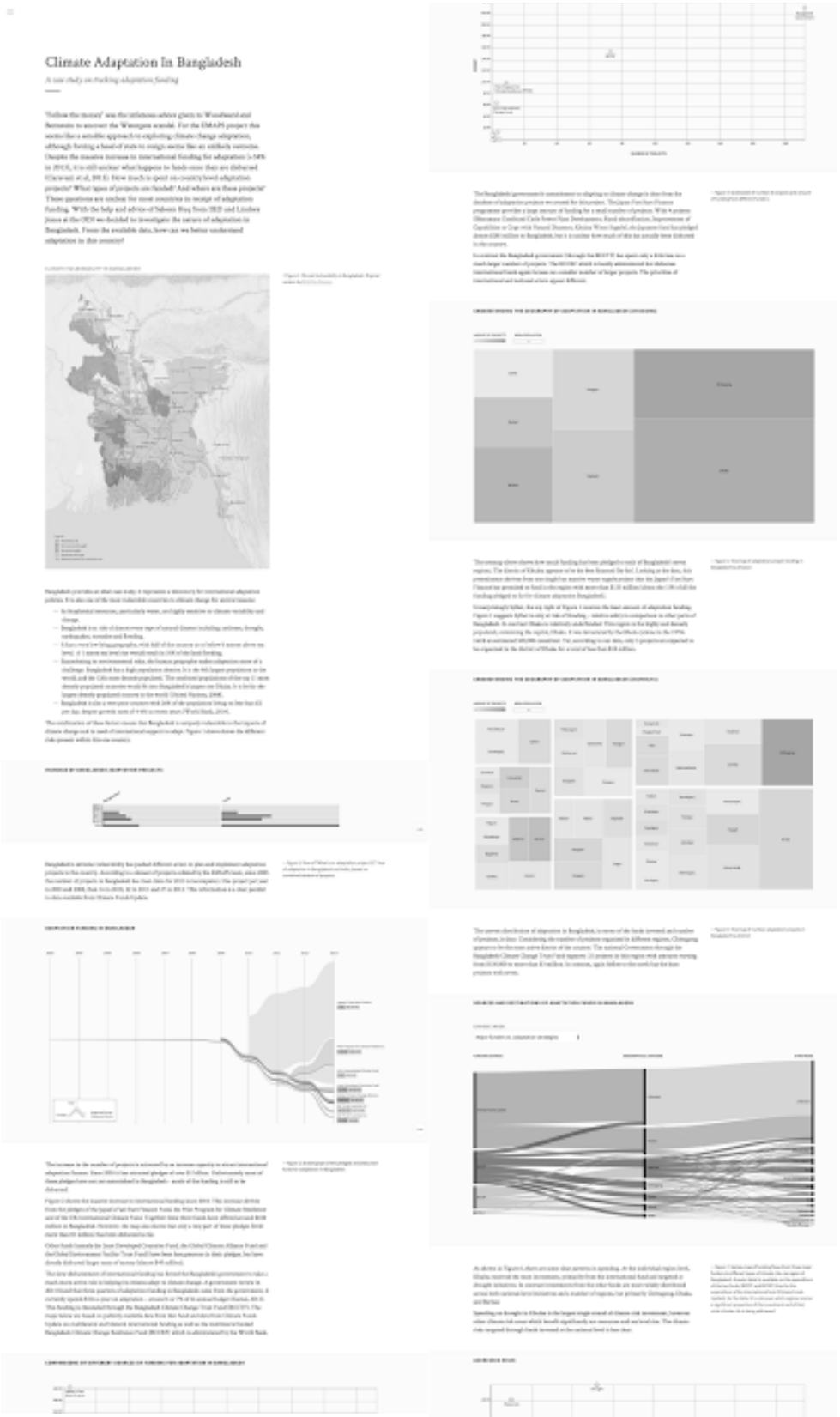
that introduce a story and allow exploration. Nevertheless, we should point to the fact that readership is not always this linear, in general it will not be. This means that most readers will not first understand the story, and then choose to explore the visualizations. We can point two main reasons for that: first, even static images have a much more immediate reception. Second, following Gibson's affordance theory (2014), that we referred to in the third chapter, the tendency is that the opposite takes place: the visual cues in interactive features will pull for the sensorimotor loops of action-reaction. So, chances are, people will probably read the titles that are visually more evident, browse through images, test the possible interactions in the interactive visualizations and only then try to understand the story.

In the procedures of access, the proposed narrative is likely to be experienced as something that is behind the presentation, and not an introduction to it. The user would have to be already engaged with the controversy and with the maps themselves to be compelled to read the story behind the presented maps. Describing the text as a narrative layer that is added to the visualizations and would be the first to be accessed is probably a description that is constructed from the point of view of the scholar who is engaged, in a sense, in the debate about the debate, in deploying the controversy. So, even when there is some involvement of the publics in the mapping process, it is important to consider that a different point of view, of the one who is actually immersed in the debate, has a specific stance, is likely to be at play. The access that is organized in layers may not work as planned because readership has for long developed non-linear procedures that mix layers. And this can be an interesting asset.

That is why we believe we should focus on a translation between the exploration on the analysis stages and in the presentation of findings in the atlas, that also shows how the procedures of the atlas are closely related to reenacting insights. Indeed, Venturini does point to a circle between exploration in analysis and the exploratory use of visualizations on the part of the public engagement: "we use exploration to produce a story, then we use it to interest people in the controversy through the exploration of the data". This is a double movement in exploration, that might also engender new cycles of mapping, thus evolving with the controversy.

The relationship between narrative and visualization is indeed a complex one, and this is especially because of all the transformations and translations between structures, and this game we have been pointing to, of creating discrete descriptions of continuities and reassembling them in new, overlaid and mapped continuities. When one builds a table out of a

narrative, they break the narrative, but can relate the resultant elements in many directions without necessarily losing track of the reference to the original narrative. When one builds a text from a work on data analysis, what are they doing? Our point is that they are not reassembling a story that came broken. In the case of controversy mapping, they are retaking the steps used to reach that graphic synthesis, and giving keys for interpreting the maps, together with a few interpretations of their own. In data journalism, they are telling the story they sourced starting from the data analysis, and using visualizations as a visual explanation.



Figures 7.6 and 7.7. Climate Adaptation in Bangladesh: Selected parts of the issue story, displaying many varieties of maps, interactive and static, and a relatively long text. Source: http://climaps.eu/#!/narrative/climate-adaptation-in-bangladesh *See Annex I, figure A1.17, for color plate.*

We believe atlas making, especially in the digital methods of social research, could gain much by carefully considering a cyclic access, that would navigate between database and narrative, visualization and text, in iterative cycles. This part seems to be sorted out quite well in many outstanding works in page design in the news: scanning and non-linear readership are already incorporated in the traditions of many outlets. On the other hand, journalistic uses of visualization could work on using the process of building the news as a guideline to organizing the narrative, like the traditions in the social sciences do so well. And, once again, we realize that scientific verifiability is in the narrative as much as in the chain of transformations of analysis. In this sense, a strong strategy to engage affected and interested publics would be to invest in turning scientific verifiability into a cyclic procedure for presentation. This procedure would amplify, at each loop, the range and richness of the controversy, evolving with it like the cycles of attentive perception described by Deleuze (1990).

8. CONCLUSIONS

During our work, we looked at the many interpretations of what information visualization is and the different inflexions of this notion in some of the many fields of knowledge and practice that use it. We described the emphasis that some fields give to the efficiency of visualizations in transmitting large amounts of information, and on its pairing with some functional descriptions of the visual cognitive processes, and argued that we could gain a better understanding of visualization and its possibilities for equipping public debate and the social production of knowledge by countering these positions with an approach that would value the collective, performative aspects of visualization. We had to trace some relations between debate, knowledge and social reality as they were organized in many different ways since the origins of modern states and inside social struggle, in order to establish visibility as a key issue to social existence and debate as a technically mediated process. Also, to discuss some effects of the contemporary attention economy and pattern recognition as a fundamental cognitive capability for societies that are heavily mediated by digital technologies.

We also had to follow a thread of mutual constitution of data and visual representations, organized by the outlining of movements of discretization and reassemblage, in order to de-naturalize some widely shared perceptions about data as a fundamental unit. That led us to explore some lines of experience, namely, time, space and context, and to outline three recurring objects, the lists, the table and the grid, that are historical and, at the same time, part of the many transformations and conversions visualizations go through in their use and circulation.

We connected these routes to fieldwork in two environments of practice where visualization plays a central role and that are fundamentally concerned with knowledge production, processes of mediation and public debate: data journalism and controversy mapping. So we had to understand how these dealt with public debate and the use of data and visualization, and how they met the challenges characteristic of their respective activities and methods. This led us to the discussion presented in the last two chapters, that connects the constitution of visualizations to the many movements in the analysis and circulation, to collective exchange, highlighting the effects of this processual development in the transformations between visualizations. This approximates analysis and presentation, and accounts for the possibility of new iterations after presentation, due to the circulation of atlases.

Now we must finish our journey by opening our scope back to the wider social processes, and establish a few issues about how visualizations are constituted into media objects and, likewise, how repurposing data in a cartographic, visual approach might develop shared spaces for public debate. We propose one last concept, which is the idea of *reverse mediation*.

For the most part of the last twenty years since the internet was broadly available, one of the reactions for issues like biases in media outlets involved, among other strategies, the creation of new independent outlets that would question hegemonic flows of information. So the web is celebrated not just for broadening access to information, but also for redistributing the tools and resources for creating, publishing and disseminating information. This collaborates for the democratization of information and forces traditional media companies to reposition their activities and procedures and to the breeding of new models, once they no longer have the exclusivity of distribution channels. Nevertheless, there is once more the problem of excess and mostly, as we have been stressing, the problem of disaggregation in the media landscape. In this sense, the fact that more channels are opened only deepens the issue: broader access to information and to publishing tools ceases to be enough. Because public debate becomes so generally integrated with technologies of information and communication, the main challenge seems to be to set common grounds for debate and relink that which composes the issues, controversies or disputes at stake in public agenda, to make them visible. Without this effort, "Social reality is out of sight": the age-old expression used by Lippman (1922) renews its relevance.

We identify an emerging field of dispute, not exactly of the origin or property of information, but of the place of mediation of the many flows of information. This dispute sprouts from a critical approach towards the traditional structures of information dissemination but also inside new models like social media – think of the algorithms that organize feeds in Facebook, for example. It does not involve specifically the creation of new and independent media outlets, but of artifacts that may enable the mapping and integration of the available flows and sources of data towards new syntheses and contextualized access.

For the purpose of our work, we have been moving around two related questions: first, can the maps produced inside controversy mapping be understood as media objects, and would this collaborate with the method's goals? Conversely, would journalism in general profit from a more cartographic approach on public agenda, that is more integrated with digital traces? We believe this contextualization or relinking of social reality through the digital traces left in

communication and information networks becomes key for public debate in today's heavily mediated societies, because it produces visible, explorable and shareable landscapes. Once data becomes such a powerful and fundamental description of many aspects of social life, visualization can become key for developing contexts for debate, much as a plane of reference is key to perception and language and to the production of knowledge that, we insist, is a social undertaking. While medias and information systems constantly propose structures and specific contextualizations to describe events and organize data flows, researchers like the ones who work with controversy mapping extract data and make maps for new interpretations, creating new mediations. We believe that data journalism might also move towards producing these new synthesis, incorporating data analysis in their routine, while developing new data-based tools to aid reporting events and their effects and repercussions. These processes revert and restructure the available flows of communication, because the traces left by actors are repurposed to generate new mediatic syntheses towards the interpretation of social reality. This leads us to advance the concept of *reverse mediation*, where social complexity is not only filtered or structured for exploration, but reprocessed. Data descriptions are extracted from informational structures, to be later turned into reassembled information, new mediatic syntheses.

But we should strengthen these claims by showing how this repurposing of communicational data has been going on, at least in the case of digital methods of social research and controversy mapping. As we have discussed, the flood of data originated from interactions in the digital networks of information and communications is seen as both an unparalleled opportunity for social analysis and, at the same time, a considerable challenge, because all this data was not produced according the needs and concerns of social research. So the powerful realization that the digital traces left by actors in social networks should be taken as evidence for studying society in general and not just online culture (ROGERS, 2009) must be countered by a careful attention to platform biases and to how society is technically mediated. We should stress the fact that these platform biases can cause distortions in the data and its interpretation, but are fundamentally part of the process of technical mediation: each interface or information architecture organizes a certain set of rules and offers some tools for enunciation, like an ethos, and overcoming or at least developing a critical stance for working with data that is bred according to these discursive codes can be quite tricky. Because of that, in a more general aspect, it will also involve understanding the many processes of mediation and encoding and decoding that go through in the circulation of messages and their

appropriation in social practices, something that will in many instances guide the process of analysis itself.

In a sense, the need for this critical understanding is not restricted to digital methods of social research: it is nowadays in the center of public debate. It happens necessarily through a critical approach of the big media companies and also the more distributed publishing tools and social media, and of their role in building public opinion or feeding public debate. This amounts in a dispute regarding the interpretation and contextualization of social events and interactions, regarding the way in which society is technically mediated.

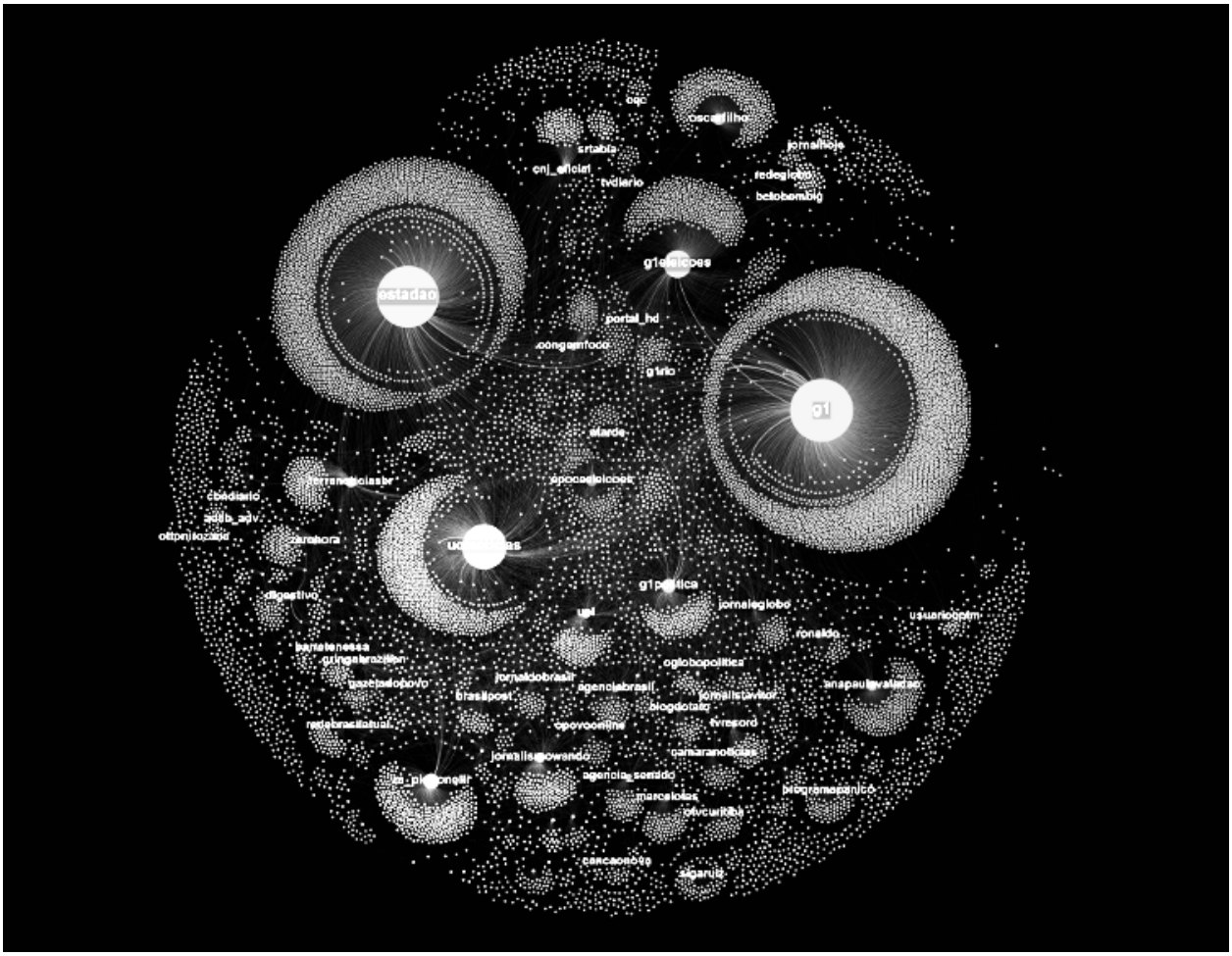


Figure 8.3. Eleições 2014. Source: http://labic.net/eleicoes2014. *See Annex I, figure A1.18, for color plate.*

In practice, this brings up many issues of how to pull apart or interpret the various platform biases, and also how to identify the strategies of powerful sources like the big medias without losing sight of the small interactions and distributed exchanges that also build the social

landscape and should be considered for non-hegemonic descriptions. It is a careful work of extracting data that is circulating in the medias and reprocessing and repurposing it in new informational forms, the visualizations. This is a challenge that is at the core of controversy mapping, but also finds some first reverberations in data journalism.

For example, in the project by Labic (figure 8.1), tweets and retweets from the day of the presidential debate in Brazil, that carried the hashtag *#debatenarecord* were mapped, showing clearly the quantitative importance of some actors like major news outlets and media personalities. This approach does show the penetration of large media companies and their encodings in social media, but much of the exchanges themselves gets masked behind the central sources, and should be revealed in order to display the richness of the controversy. Also, there is no hint of how the interpretation of these dominant encodings are being decoded and incorporated into more distributed discursive practices.

The project Emaps, on the other hand, faces this challenge from a different perspective: they present a profile of Twitter accounts (figure 8.2) made at the DMI Summer School 2013, where the emphasis was in revealing how the subject of climate change was made into a matter of concern to be shared, that is, how it was framed and how the correspondent discursive features were built in Twitter. For that, they captured tweets from november 2012 to may 2013 that contained either hashtags or keywords in their content that were related to climate change. From this corpus, they defined four major climate change issues, queried by the terms [skeptic], [mitigation], [adaptation], [conflict OR violence], in order to generate some metrics for profiling. The graphic shows the popularity of each issue and informations like the main related hashtags and how they compare across issues. It also shows the main sources linked and the most active profiles, which gives us interesting information on the subject. The main difference here is that hashtags are not taken by face value: all these correlations aid in developing a better understanding of actually what they mean in the space they assemble. The emphasis stays on the debate and how it is shared on Twitter.

Figure 8.2. Profiling Adaptation And Its Place In Climate Change Debates With Twitter (I). Source: http://climaps.eu/#!/map/profiling-adaptation-and-its-place-in-climate-change-debates-with-twitter-i

On the results of the Emaps project, on the Climaps.eu site, this graphic is associated to a map (figure 8.3), that displays the main hashtags found and their proximity according to co-occurence in tweets, inside the set of tweets that contained the hashtags or keywords “climate change” or “global warming” between November 2012 and November 2013. This gives us yet another point of access to examining the debate on Twitter, that helps us understanding how different hashtags may be related, and thus translated in their use. The chosen theme (climate change) organizes a Twitter *space*, a landscape to be explored for its specific issues. Both visualizations highlight hashtags and keywords as powerful traces for organizing maps of debates, and, as their positions may shift and get reorganized, so do the many actors and positions involved. This comparative and relational approach also enriches the discussion because it favors a critical approach on the hashtags as platform-specific resources, in order to take a step back and evaluate also their effects and biases. This work reveals a more relational, qualitative network of arguments and gives us material to discuss the processes of attribution of signs in different maps of meaning, that is involved in every decoding process.

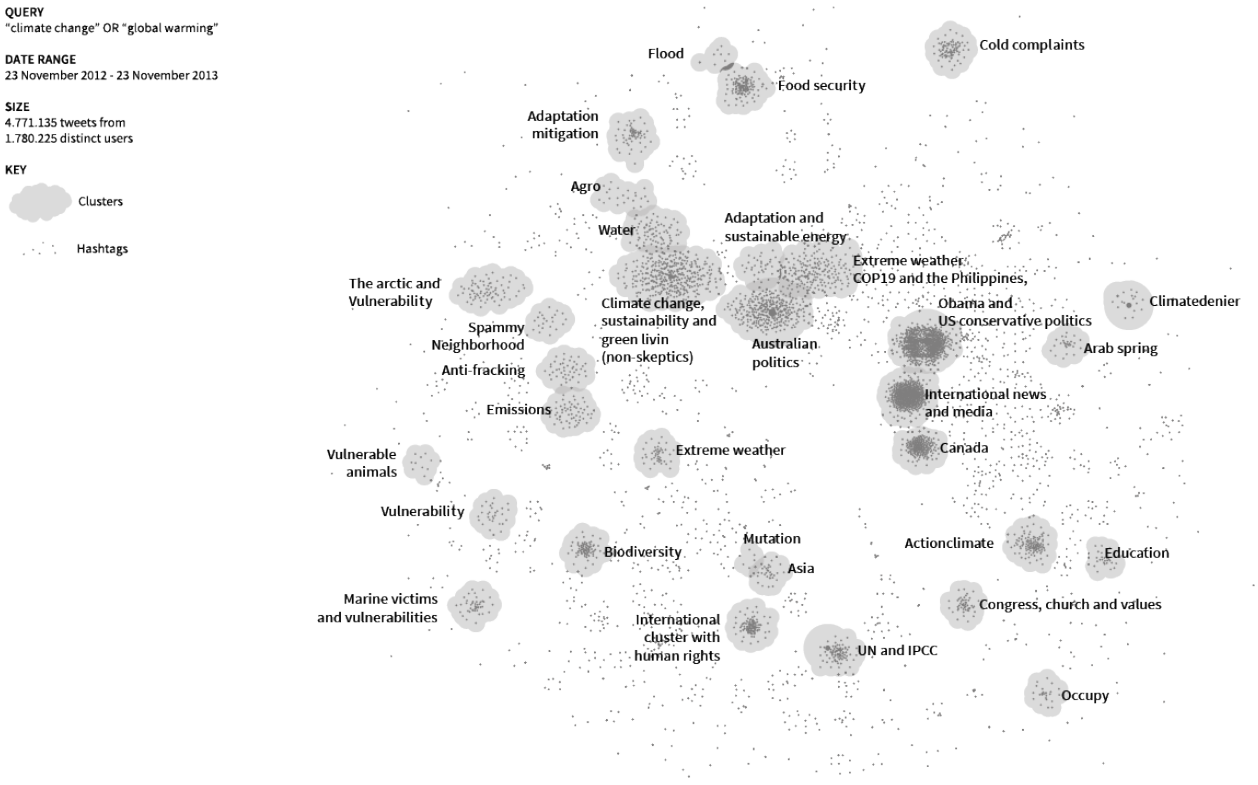


Figure 8.2. Profiling Adaptation And Its Place In Climate Change Debates With Twitter (I). Source: http://climaps.eu/#!/map/profiling-adaptation-and-its-place-in-climate-change-debates-with-twitter-ii

Until now, we have been giving much emphasis to the challenge of assembling useful representations, visual contexts, out of a scattered landscape, which, in the third chapter, we

related to the process of discretization in the technologies for recording and displaying information. Nevertheless, as we have proposed in the second chapter, the problem of a scattered info-communicational landscape is also connected to the formation, in any society, of distinct discursive spaces, that, according to Hall, will reflect the continuation, reinforcement or challenging of power and ideology inside discursive practices. We would like to advance that the constitution of discursive spaces has its effects on the communication infrastructures and interfaces, that also constitute frames for what can be said and how it should be said: they participate in structuring discourse and information through many different matrixes, in general relying on different codes, inside varied maps of meaning.

This can be considered yet another kind of dominance at the encoding moment, but this time it is brought on by the communicational infrastructure itself: the effects of the information architectures and the many features for managing the information flows. Again, a strong example comes from the social media: sites like Facebook, for example, are constantly making improvements in their interface's structure and adding new features. For a while along its first years, there were many of such updates, and at each one we had the impression that it would be hard to adapt, but soon found ourselves not having a clear idea of in what consisted the changes themselves. The features had quickly faded into the background, as an encoded base structure that maps the means for our exchanges and debates. So with online media we have both complementary aspects: more visibility and more structuring to canalize discourse and debate into the informational flow. There is a stage of dispute of meaning attributions that is displayed in how the computer interfaces are organized, how data is visualized, how the places of mediation create paths and structures and participate in broader production of social knowledge. In other words, the compositions derived from the architecture of information become more spread in the consumption moments but at the same time more varied and naturalized. We believe that most of the concern for the effects of the platform in the digital methods of social research comes from this entangled relation between public debate and dominant information architectures, that are fundamental for professional outlets as well as for individual users of social media or self-publishing platforms.

While traditional print and broadcast media had almost exclusivity on publishing channels, it might have been adequate, even though not precise, to describe the production and circulation of messages in society through, on the one side, major channels and codes, and, on the other side either individual decoding, mainly as a psychological behavior, or statistical behavior that could be extrapolated into conclusions about collective message reception. Substantially,

the sphere of public debate as a social phenomena was seen as an untrackable mess. These approaches, even though they do bring much insight to the table, miss the point that communication has always been technically mediated, in every level, and that, in that sense, mediations are multidirectional. With the spread of online platforms it becomes unavoidable to carefully consider other distributed mediations as part of the media landscape, but there is no actual modification of communication flows, just intensification of exchanges.

The work conducted in digital social research in general, and in the journalistic practices that incorporate these methods or develop combinations with traditional models, is creating new mediations, new interfaces in the form of maps, that repurpose information flows. Once there is a general recognition that technical mediation produces traces that can be used for describing society in general, these methods are discovering the strength of extracting data from information to reassess information: this translates many encodings, many structures, many discursive spaces into an assembled landscape. We understand that this reprocessing and repurposing of information reverses mediations, not in the sense of inverting flows, but in the sense of problematizing and surpassing information structures. After all architecture, there might be a recovery of the plasticity for information through data extraction and visual analysis.

We stress terms like “reversing” or “repurposing” in order to avoid thinking of visualizations, in the context of reverse mediations, as yet another layer that piles up our already *overlayered* experiences. We have found that, when one seeks to map debates, either in social sciences or in journalism, there is the risk that the representations start to be seen mostly as meta-debates, the debate about the debate, accumulating indefinitely layers of representation of debates about debates. Indeed, even without mapping, debates will of course take place with the use of many and contradictory representations of a certain issue, and there would be neither an issue nor an engaged public if all the representations and codes coincided. So we believe this movement of reversing, or bringing the informational structures inside out, is crucial for equipping citizenship in heavily mediated societies, not exactly because of what is represented, but because of the many points of access it offers to a complex communicational landscape. The debate about the debate is part of the work of social scientists and of journalists, but will not necessarily be a procedure in public debate, especially considering the important communicational moment of decoding messages and incorporating them as actions, into social practices. Venturini (2012) advances that controversy mapping is not a method for social intervention, which is a very innovative and powerful claim, but, on the other hand, we

should keep in mind that the adoption of those maps by the engaged publics must include social intervention, but in the sense of political action inside the controversy. And we are using the term *political* in its broader sense, that runs across states, institutions and everyday life. Like Virno explains, it is with the publicization of intellect that political action gets intertwined with the basic features of language and cognition. When visualizations that aim at creating shared landscapes are appropriated by the concerned publics, one should expect, and aim for, a coincidence between debate and action.

REFERENCES

ARRUDA, Magali. In: DANTAS, Marcello. (Org.) **100 objetos: Peter Greenaway**. Rio de Janeiro: CCBB, 1998.

AUSTIN, John; SABISÀ, Marina (ed.). *How to do things with words*. Harvard: Harvard University Press, 1975.

BARABÁSI, Albert-László. **Network science**. 2014. Available: http://barabasi.com/networksciencebook. Access: 10/03/2015.

BAUDRILLARD, Jean. **Simulacros e Simulações.** Lisboa: Relógio D´Água, 1991.

BEDERSON, Benjamin B. (org.); SCHNEIDERMAN, Ben (org.). **The craft of information visualization: readings and refletions**. San Francisco: Morgan Kaufmann, 2003.

BELTING, Hans. **Antropología de la imagen**. Madrid, Katz Editores: 2010.

BENEDIKT, Michael. Information in space is space in information. In: MICHELSEN, Anders; STIERNFELT, Frederik (orgs.). **Images from Afar: Scientific Visualization – an Anthology**. Akademisk Forlag: Copenhagen, 1996.

BENEDIKT, Michael (org). **Cyberspace: first steps**. Cambridge: MIT Press, 1994.

BERGER, P.; LUCKMANN, T. **A Construção Social da Realidade**. Petropólis, Editora Vozes, 2008.

BERGSON, Henri. **Matéria e memória: ensaio sobre a relação do corpo com o espírito**. Martins Fontes: São Paulo, 1999.

_____. **A evolução criadora**. Martins Fontes: São Paulo, 1995.

BERTIN, Jacques. **Semiology of graphics: diagrams, networks, maps**. ESRI Press: New York, 2011.

BORGES, Jorge Luis. **La biblioteca de Babel**. Buenos Aires: Editora Argentina, 2001.

BÖRNER, Katy; POLLEY, David E. **Visual insights: a pactical guide to making sense of data**. MIT Press: London, 2014.

BOTTÉRO, Jean. A escrita e a formação da inteligência na Mesopotâmia Antiga. In: BOTTÉRO, Jean, MORRISON, Ken et al. **Cultura, pensamento e escrita**. São Paulo: Ed Ática, 1995. p. 9-46.

BOULLIER, Dominique. Composition médiatique d'un monde commun à partir du pluralisme des régimes d'attention. In CHARDEL, Pierre Antoine; GOSSART, Cédric; REBER, Bernard. **Conflit des interprétations dans la société de l'information - Ethiques e politiques de l'environnement**. Hermes Editions: Paris, 2012.

_____. Le web immersif. In: **Quaderni**. N. 66, Printemps 2008. Cybersp@ce & Territories. p. 67-80.

CAIRO, Alberto. **The Functional Art: An introduction to information graphics and visualization**. San Francisco: New Riders, 2012.

CAPURRO, Rafael. Pasado, presente y futuro de la noción de información. *Logeion: Filosofia da Informação* 1.1, 2014, p.108-133.

CARD, Stuart K.; MACKINLAY, Jock; SCHNELDERMAN, Ben. **Readings in Information Visualization: Using Vision to Think (Interactive Technologies)**. San Francisco, Morgan Kaufmann, 1999.

CARDOSO, Rafael Denis. **Design para um mundo complexo**. São Paulo, Cosac & Naify, 2012.

CHARTIER, Roger. **A aventura do livro**. São Paulo: Fundação Editora UNESP, 1998.

CHION, Michel. **Le son au cinéma**. Paris: Cahiers du Cinéma, 1985.

CITTON, Yves. Le marketing entre éonomie de l'attention et explotation culturelle. In BOURGNE, Patrick (dir.). **Le marketing, poison ou remède? Les effets du marketing dans une societé en crise**. Cormelles-Le-Royal: Editions EMS, 2013, p.179-199.

COUZINET, V. Des pratiques érudites à la recherche: bibliographie, bibiologie. In: GARDIÈS, C. **Approche de l'information-documentation: concepts fondateurs**. Toulouse: Cédaduès-Éditions, 2011. p. 167-186.

CRAMER, J.; McDEVITT, M. Ethnographic journalism. **Qualitative Research in Journalism: Taking it to the Streets**, Iorio, S. (ed.), Lawrence Erlbaum, Mahwah, NJ, 127-144, 2004.

CRARY, Jonathan. A visão que se desprende: Manet e o observador atento do século XIX. In: CHARNEY, Leo; SCHWARTZ, Vanessa R.; THOMPSON, Regina. **O cinema e a invenção da vida moderna**. São Paulo: Cosac & Naify, 2007.

_____. **Techniques of the observer: on vision and modernity in the nineteenth century**. October Books, MIT Press: Cambridge, Massachusetts, 1990.

DASTON, Lorraine; GALISON, Peter. **Objectivity**. Zone Books: Brooklyn, 2007.

DELEUZE, Gilles. **Francis Bacon: the logic of sensation**. Continuum: London, 2003.

_____. **Bergsonismo**. Editora 34: São Paulo, 1999.

_____. **A imagem-tempo**. Editora São Paulo: Brasiliense, 1990.

_____. **Cinema1: A imagem-movimento**. São Paulo: Brasiliense, 1985.

DELEUZE, Gilles; GUATTARI, Felix. **Mil platôs: capitalismo e esquizofrenia**, vol.1. Editora 34: São Paulo, 1995.

DEWEY, John; ROGERS, Melvin L (ed). **The Public and its Problems: an Essay in Political Inquiry**. Pennsylvania: Penn State University Press, 2012.

DISALVO, Carl. **Adversarial Design**. MIT Press: Massachussets, 2012.

DURHAM, M. G. (1998), On the relevance of standpoint epistemology to the practice of journalism: the case for 'strong objectivity'. **Communication Theory** Eight: two. May 1998, p.117-140.

EISENSTEIN, Sergei. **A forma do filme**. Rio de Janeiro: Zahar, 2002.

FLUSSER, Vilém. **O mundo codificado**. São Paulo: Cosac Naify, 2007.

FONSECA, Edson Nery da. A bibliografia como ciência: da crítica textual à Bibliometria. **Revista Brasileira de Biblioteconomia e Documentação**, Brasília, DF, v. 12, n. 1-2, p. [29]-38, jan./jun. 1979.

FOUCAULT, Michel. O que são as Luzes? In FOUCAULT, Michel. **Ditos e escritos**, v.II. Rio de Janeiro: Forense Universitária, 2009.

FOUCAULT, Michel. Outros Espaços. In: **Ditos e escritos**, v.IV. Rio de Janeiro, Forense Universitária, 2003.

FOUCAULT, Michel. **História da sexualidade I: a vontade de saber**. 7.ed. Rio de Janeiro: Graal, 1985.

FRY, Benjamin Jothe. **Computational Information Design**. 2004, 170f. Thesis (PhD in Media Arts and Sciences). School of Architecture and Planning, Massachusetts Institute of Technology, Massachusetts, 2004.

GEERTZ, C. **The Interpretation of Cultures**. New York, Basic Books, Inc. Publishers, 1973.

GIBSON, James. The theory of affordances. In: GIESEKING, Jen J. *et alii* (ed.). **The people, place and space reader**. Routledge: London, 2014.

GLEICK, James. **The information: a history, a theory, a flood**. New York: Pantheon Books, 2011.

GOMES, Wilson. Esfera pública política e media: com Habermas, contra Habermas. In GOMES, Wilson; MAIA, Rousuley C. M. **Comunicação e democracia: problemas e perspectivas**. Rio de Janeiro: Paulus, 2005.

GOODY, Jack. **La raison graphique: la domestication de la pensée sauvage**. Lonrai, Les Edition de Minuit, 1997.

GRAY, J. (ed.); BOUNEGRU, L. (ed.); CHAMBERS, L. (ed.). **The Data Journalism Handbook**, 1.0 beta. http://datajournalismhandbook.org/1.0/en/ (Access at 13/06/2014)

HABERMAS, Jürgen. **The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society** (Studies in Contemporary German Social Thought). MIT Press, Cambridge, MA, 1991.

HACKING, Ian. The looping effects of human kinds, In SPERBER, Dan (ed.); PREMACK, David (ed.); PREMACK, Ann James (ed.). **Causal cognition: a multidisciplinary debate**. New York: Oxford University Press, 1995.

HALEY, J.B.; WOODWARD, David. The History of Cartography, vol.2, book 2. Chicago: University of Chicago Press, 1994.

_____. The History of Cartography, vol.2, book 1. Chicago: University of Chicago Press, 1992.

_____. The History of Cartography, vol.1. Chicago: University of Chicago Press, 1987.

HALL, Stuart. Encoding, decoding. In DURING, Simon (ed.). **The cultural studies reader** 4. Routledge: London,1993, p.80-103.

HEALY, Kieran. MOODY, James. Data visualization in sociology. **Annual review of sociology** 40, 2014, p. 105-128.

HOGAN, Adam. **Visualizing the history of philosophy as a social network: the problem with Hegel**. Available: http://www.designandanalytics.com/visualizing-the-history-of-philosophy-as-a-social-network-the-problem-with-hegel Access: 2012/09/20

HORCH, Rosemarie Erika. Bibliografia Textual. **Revista Brasileira de Biblioteconomia e Documentação**, Brasília, DF, v. 11, n. 3-4, p. 147-154, jul./dez. 1978.

JOHNSON, Steven. **Emergence**: the connected lives of brains, cities and software. Scribner: Montreal, 2004.

_____. **Cultura da interface**: como o computador transforma nossa maneira de criar e comunicar. Rio de Janeiro: Jorge Zahar, 2001.

KANNO, Mario. **Infografe**: Como e porque usar infográficos para criar visualizações e comunicar de forma imediata e eficiente. São Paulo: Infolide, 2013.

KANT, Immanuel. Immanuel **Kant**: textos seletos. Petrópolis: Vozes, 2005.

KASTRUP, Virginia. **A invenção de si e do mundo**: uma introdução do tempo e do coletivo no estudo da cognição. Editora Autêntica: Belo Horizonte, 2007.

KEIM, Daniel A. Information visualization and visual data mining. **IEEE transactions on visualization and computergraphics**, vol.7, no.1, January-March 2002.

KIRSH, David. Thinking with External Representations. **AI and Society**. Vol.25, p.441–454. Springer: London, 2010a.

_____. *Explaining Artifact Evolution*. In Malafouris, L. and Renfrew, C. (eds.) **The Cognitive Life of Things: Recasting the boundaries of the mind**. McDonald Institute for Archaeological Research, 2010b.

_____. A Few Thoughts on Cognitive Overload, **Intellectica**, p.19-51, 2000a.

_____. Distributed Cognition, A Methodological Note. **Pragmatics & Cognition**, ed. 14, vol. 2, p. 249-262, 2000b.

_____; MAGLIO, P. On Distinguishing Epistemic from Pragmatic Actions. In: **Journal of Cognitive Science**. Cognitive Science Society: Austin,1995.

LATOUR, Bruno (org.). Vues de l'esprit, **Culture & Technique** no.14. Centre de Recherche sur la Culture Technique: Neuilly-sur-Seine,1985. Available: http://documents.irevues.inist.fr/handle/2042/29797 (Access: 11/09/2013)

_____; BIEZUNSKI, Michel. **La science en action: introduction à la sociologie des sciences**. La Découverte, Paris, 2005.

_____. **Reassembling the social: an introduction to actor-network theory**. Oxford University Press: Oxford, 2005.

_____. On TechnicaI Mediation - Philosophy, Sociology, Genealogy. in **Common Knowledge**, Vol.3, N.2, p.29-64, Fall 1994.

_____. et al. The Whole is Always Smaller Than Its Parts A Digital Test of Gabriel Tarde's Monads. **British Journal of Sociology**, 2009.

_____. **Pandora's hope: essays on the reality of science studies**. Harvard: Harvard University Press, 1999.

LAW, J. **Notes on the theory of actor network: ordering, strategy and heterogeneity**. Lancaster: Centre for Science Studies, Lancaster University, 1992. http://www.comp.lancs.ac.uk/sociology/papers/Law-Notes-on-ANT.pdf

LÉVY, Pierre. **Tecnologias da inteligência**. Ed. 34: São Paulo, 1997.

_____. **A ideografia dinâmica: rumo a uma imaginação artificial?** Edições Loyola: São Paulo, 1998.

LIPOVETSKY, Gilles. **O Império do efêmero: a moda e seu destino nas sociedades modernas**. São Paulo: Companhia das Letras, 2007.

LIPPMANN, Walter. **Public Opinion**. Free Press, Nova Iorque, 1922.

LYNCH, Michael. Le retine exteriorisé. In: LATOUR, Bruno (org.). **Vues de l'esprit, Culture & Technique** no.14. Centre de Recherche sur la Culture Technique: Neuilly-sur-Seine,1985.

MANOVICH, Lev. Database, datastream, timeline: the forms of social media (part1). **Software Studies Initiative Blog**. October 27th 2012. Available: http://lab.softwarestudies.com/2012/10/data-stream-database-timeline-new.html (Access: 11/12/2014)

_____. What is visualization? **paj: The Journal of the Initiative for Digital Humanities, Media, and Culture** 2.1, 2010. Available: https://journals.tdl.org/paj/index.php/paj/article/view/19 (Access: 11/12/2014).

_____. **The language of new media**. Cambridge: MIT Press, 2001.

_____. **Database as a Symbolic Form**. Cambridge: MIT Press 1998. Available: http://transcriptions.english.ucsb.edu/archive/courses/warner/english197/Schedule_files/Manovich/Database_as_symbolic_form.htm (Access: 11/12/2014)

MARRES, Noortje. **Why Map Issues? How digitization is changing the settings of controversy analysis and what to do about it**. (forthcoming)

_____. Issues spark a public into being. A key but often forgotten point of the Lippmann-Dewey debate. Em Latour, Bruno; Weibel, Peter (eds.). **Making Things Public**, MIT Press,Cambridge MA, 2005

MARRES, N.; Weltevrede, E. Scraping the Social? Issues in realtime social research. **Journal of Cultural Economy**, 2012.

MASSARANI, Luisa. **A divulgação científica no Rio de Janeiro: Algumas reflexões sobre a década de 20**. Orientadores: Lena Vânia Ribeiro Pinheiro, Dr. Ildeu de Castro Moreira, Dr.. IBICT: Rio de Janeiro, 1998. Dissertação (Mestrado em Tecnologia da Informação). Disponível em http://www.cciencia.ufrj.br/publicacoes/dissertacoes/Massarani_tese.PDF (Acesso em 12/09/2013).

MATOS, Nuria E. Pérez. La bibliografía, bibliometría y las ciencias afines. **ACIMED** v.10 n.3 Ciudad de La Habana Mayo.-jun. 2002.

MAYER, K. Scientific images? How touching! **Science, Technology & Innovation Studies**, 7(1), p. 29 – 45, 2011

MAYNARD, Patrick. **Drawing distinctions: the varieties of graphic expression**. Cornell University Press: Ithaca, 2005.

McNAIR, B. **The sociology of journalism**. New York: Bloomsbury Academic, 1998.

MEYER, P. **The New Precision Journalism**. 1991. http://www.unc.edu/~pmeyer/book/

MULTHAUF, Robert P. The Introduction of Self-Registering Meteorological Instruments. **United States National Museum Bulletin**, n.228. Washington: Smithsonian Institution, 1961. p. 95-116. Available at http://www.gutenberg.org/files/32482/32482-h/32482-h.htm (Acesso em 24/01/2015).

MUNZNER, Tamara. Process and Pitfalls in Writing Information Visualization Research Papers. Em KERREN, Andreas; STASKO, John T.; FEKETE, Jean-Daniel; NORTH, Chris (editores). **Information Visualization: Human-Centered Issues and Perspectives**. Springer LNCS Volume 4950, p 134-153, 2008. Disponível em http://www.cs.ubc.ca/labs/imager/tr/2008/pitfalls (Acesso em 12/09/2013)

NOVELLO, Mário. **O que é cosmologia? A revolução do pensamento cosmológico**. Jorge Zahar Editore: Rio de Janeiro, 2006.

NOVEMBER, V.; CAMACHO-HÜBNER, E.; LATOUR, B. 2010. **Entering a RiskyTerritory: Space in the Age of Digital Navigation**. <http://www.bruno-latour.fr/sites/default/files/117-MAP-DIGITAL-GB.pdf> (acesso em 01/05/2012)

OFFENHUBER, D. Visual Anecdote. **Leonardo**, 43(4), 2010, p.367–374. Available: http://dspace.mit.edu/handle/1721.1/60338 (Access: 01/05/2014)

PANOFSKY, Erwin; WOOD, Christopher S. **Perspective as a symbolic form**. Zone Books: New York, 1991.

PARENTE, André. **O virtual e o hipertextual**. Editora Pazulin: Rio de Janeiro, 1999.

_____. **Imagem-máquina: a era das tecnologias do virtual**. Ed. 34: Rio de Janeiro, 1993.

PETERSON, M. A. Journalism as trope. **Anthropology News**, April. p. 8-9, 2010.

PLAYFAIR, William. **Playfair's Commercial and Political Atlas and Statistical Breviary**. Cambridge: Cambridge University Press, 2005.

POPPER, Karl. **The logic of scientific discovery**. London: Routledge, 2002.

RANCIÉRE, Jacques. **O desentendimento**. Ed. 34: São Paulo, 1996.

ROGERS, Richard. **The end of the virtual**: digital methods. Cambridge: MIT Press, 2013.

ROQUE, Tatiana. **História da matemática**: uma visão crítica, desfazendo mitos e lendas. Zahar Editores: Rio de Janeiro, 2012.

ROSENBERG, Daniel; GRAFTON, Anthony. **Cartographies of time**: a history of the timeline. Princeton Architectural Press: London, 2010.

SERRES, Michel. **Les origines de la géométrie**. Champs Sciences: Paris, 2004.

_____. **Hominiscências:** o começo de uma outra humanidade? Bertrand Brasil: Rio de Janeiro, 2003.

_____; FAROUKI, Nayla. **Paysages des Sciences**. Paris: Éditions Le Pommier-Fayard, 1999.

SHANNON, Claude E. A Mathematical Theory of Communication. **The Bell System Technical Journal**, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

SLOTERDIJK, Peter. **Spheres, vol.I - Bubbles: Microspherology**. Los Angeles: Semiotext(e), 2011.

SLOTERDIJK, Peter. **Regras para o parque humano**: uma resposta à carta de Heidegger sobre o humanismo. Editora Estação Liberdade: São Paulo, 2001.

TARDE, Gabriel. L'archéologie et la statistique. **Revue philosophique**, tome XVI, 1883. P. 363-384 et 492- 511.

TUFTE, Edward R. **The Visual Display of Quantitative Information.** Cheshire: Graphics Press, 1983.

_____. **Visual Explanations**: Images and Quantities, Evidence and Narrative. Connecticut: Graphics Press, 1997.

VARELA, Francisco. **Invitation aux sciences cognitives**. Éditions du Seuil: Paris, 1989.

VENTURINI, Tommaso. Diving in magma: How to explore controversies with actor-network theory. **Public Understanding of Science**, 2009.

_____. et al. **Designing Controversies and their Publics**. (*forthcomming*)

_____; LATOUR, Bruno. The Social Fabric: Digital Traces and Quali-quantitative Methods. **Proceedings of Future En Seine**, p. 15-30, 2010. Available: http://www.medialab.sciences-po.fr/publications/Venturini_Latour-The_Social_Fabric.pdf . Access: 05/032015.

VIRILIO, Paul. **O Espaço crítico**. Tradução Paulo Roberto Pires. 1. ed. Rio de Janeiro: Editora 34, 1993.

VIRNO, Paolo. **Virtuosismo e revolução**. Rio de Janeiro: Civilização Brasileira, 2008.

WARE, Colin. **Information visualization: perception for design**. San Francisco: Morgan Kaufmann, 2003.

WILLIAMS, Henry Smith. **A history of science**. New York: Harper Collins, 1904.

WINKLER, Capt. On sea charts formerly used in the Marshall Islands, with notices on the navigation of these islanders in general. **Smithsonian Institute Report for 1899**, 54, 487-508, 1901. Disponível em: http://www.ethnomath.org/resources/winkler1901.pdf (Acesso em 11/09/2013)

WOLF, Mauro. **Teorias da comunicação**. Lisboa, Presença, 1987.

WURMAN, Saul. **Information anxiety 2**. Indianapolis: Que, 2000.

ZER-AVIV, Mushon. Disinformation Visualization: How to lie with datavis. In **Visualizing Information for Advocacy**, January 31, 2014. Available: https://visualisingadvocacy.org/blog/disinformation-visualization-how-lie-datavis (Access: 09/02/2015)

ANNEX A – COLOR PLATES

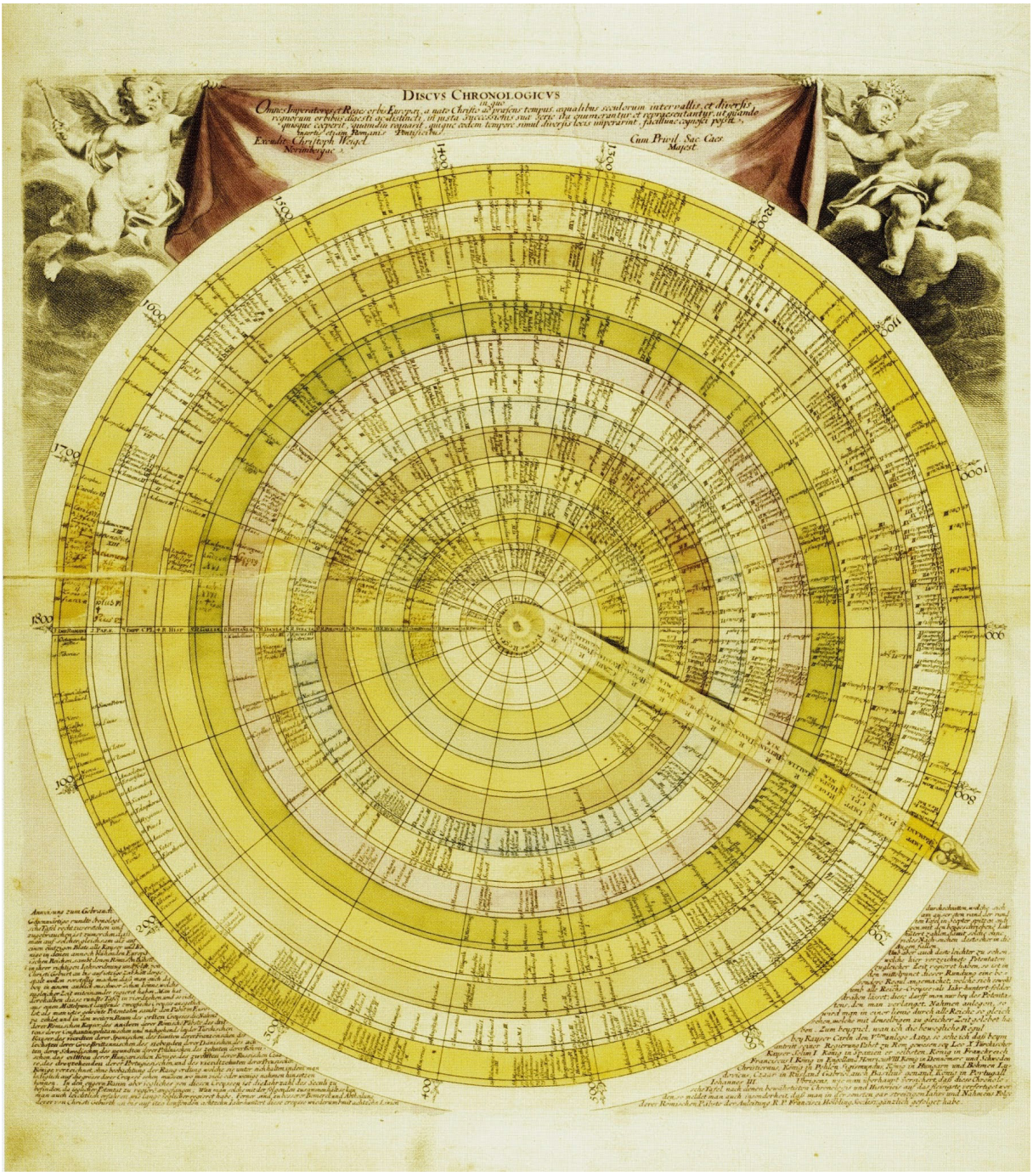


Figure A1.2. Discus Chronologicus, published in the early 1820s by Christoph Wiegel.
Source: https://www.flickr.com/photos/bibliodyssey/4501667579/sizes/o/

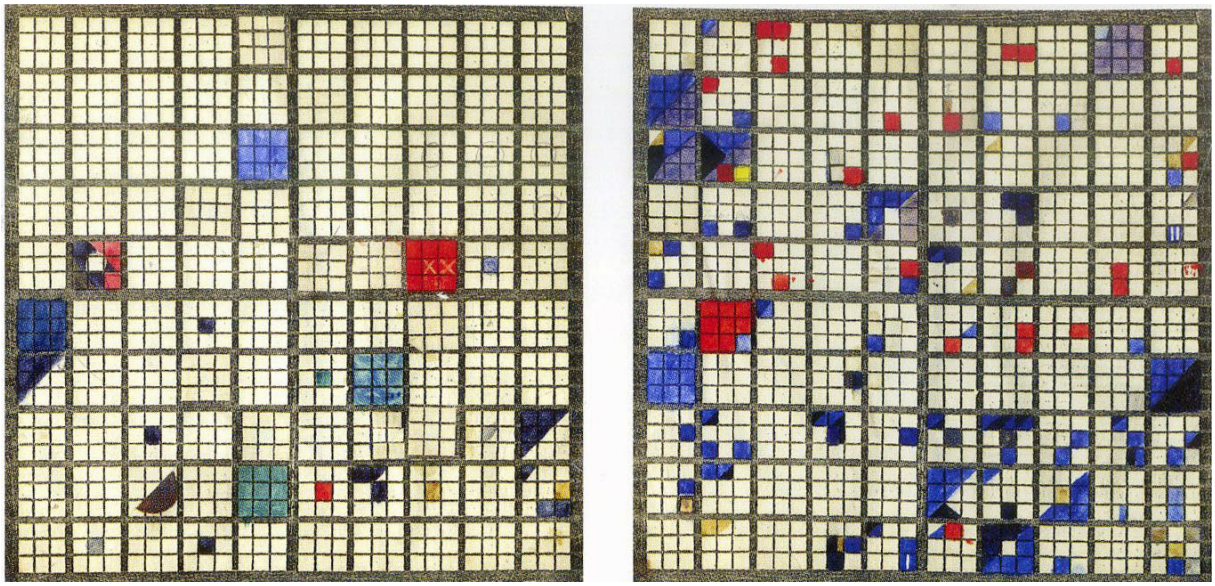


Figure A1.3. Elizabeth Palmer Peabody, Polish-American System of Chronology, 1850. Source: Rosemberg; Grafton: 2010, p.205.

Figure A1.4. Photography with 360° fisheye lens, by Randy Scott Slavin / Rex Feature
Source: http://www.dailymail.co.uk/news/article-2131638/Who-needs-Instagram-Photographer-uses-traditional-fisheye-technique-capture-extraordinary-landscapes-America.html

Figure A1.5. *Las Meninas* ou *La familia de Felipe IV*, pintura de Diego Velázquez (1656)
Source: http://commons.wikimedia.org/wiki/File:Las_Meninas_(1656),_by_Velazquez.jpg

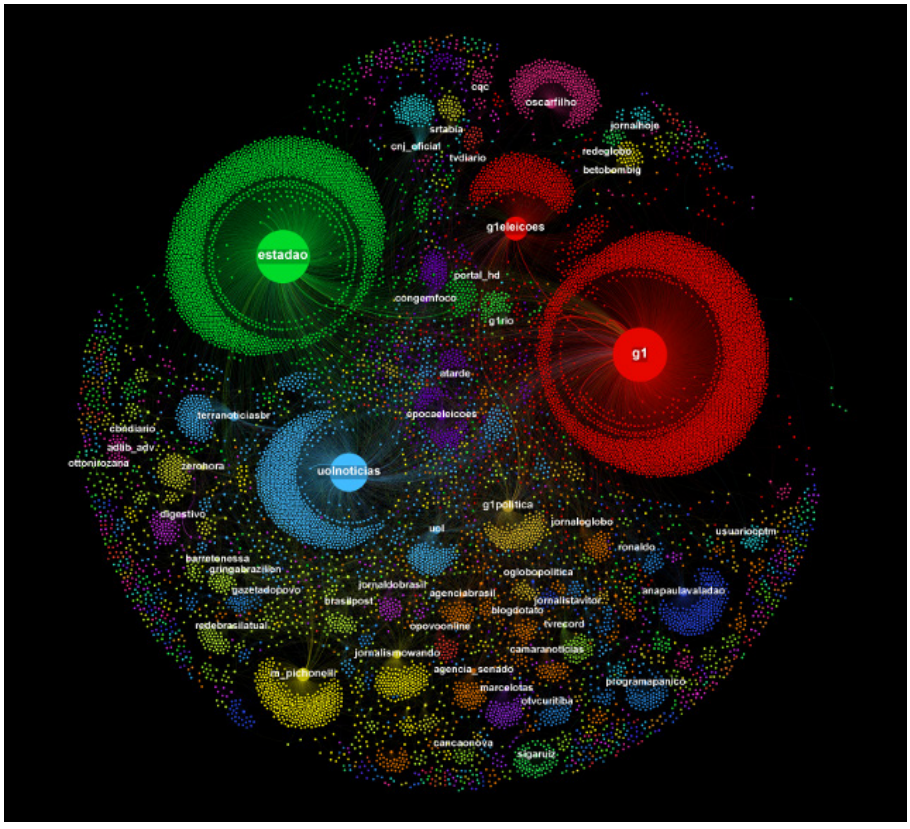


Figure A1.19. Eleições 2014. Source: http://lab-ic.net/eleicoes2014

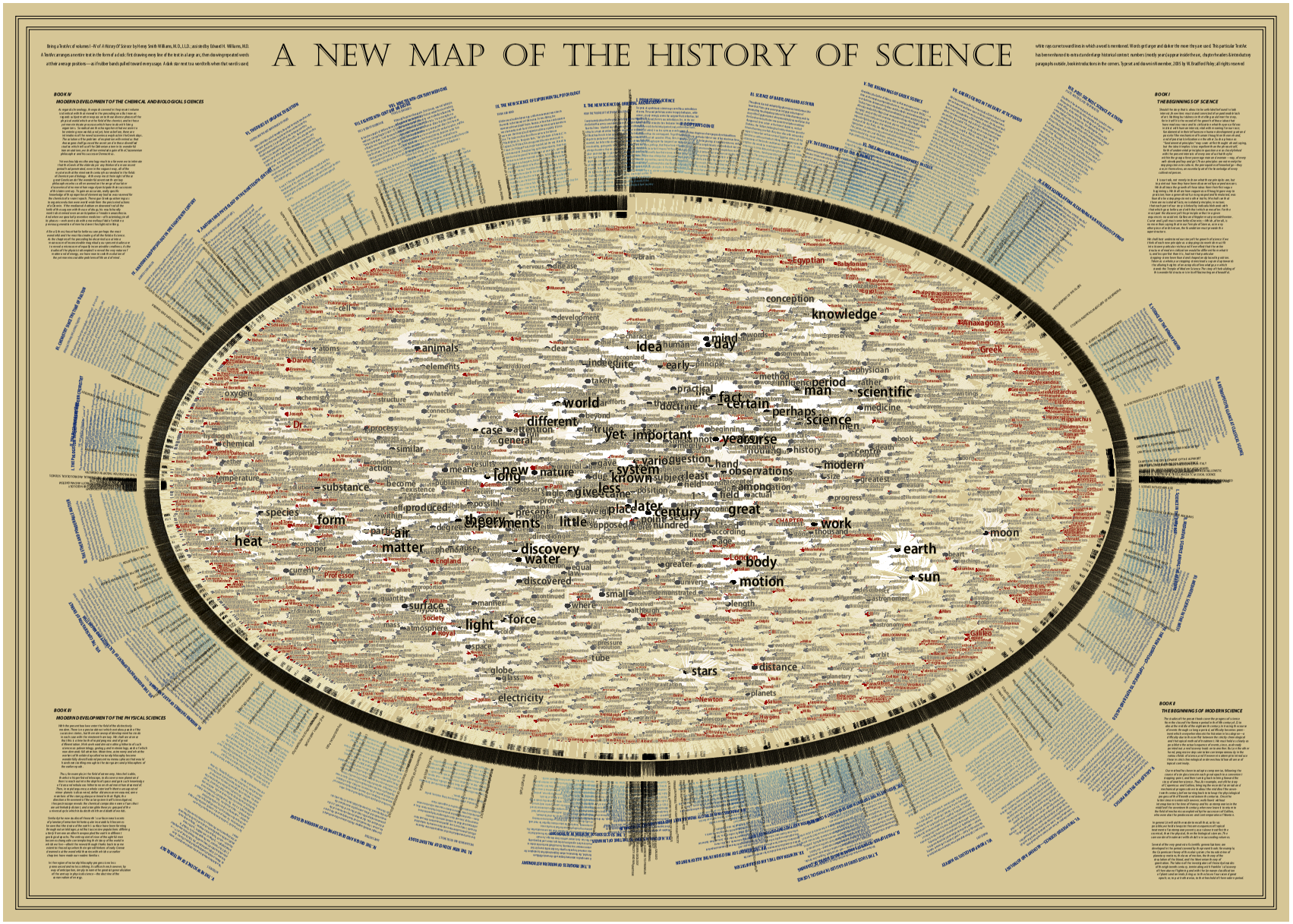


Figure A1.6: *TextArc visualization of The History of Science* (2006), by Bradford Paley

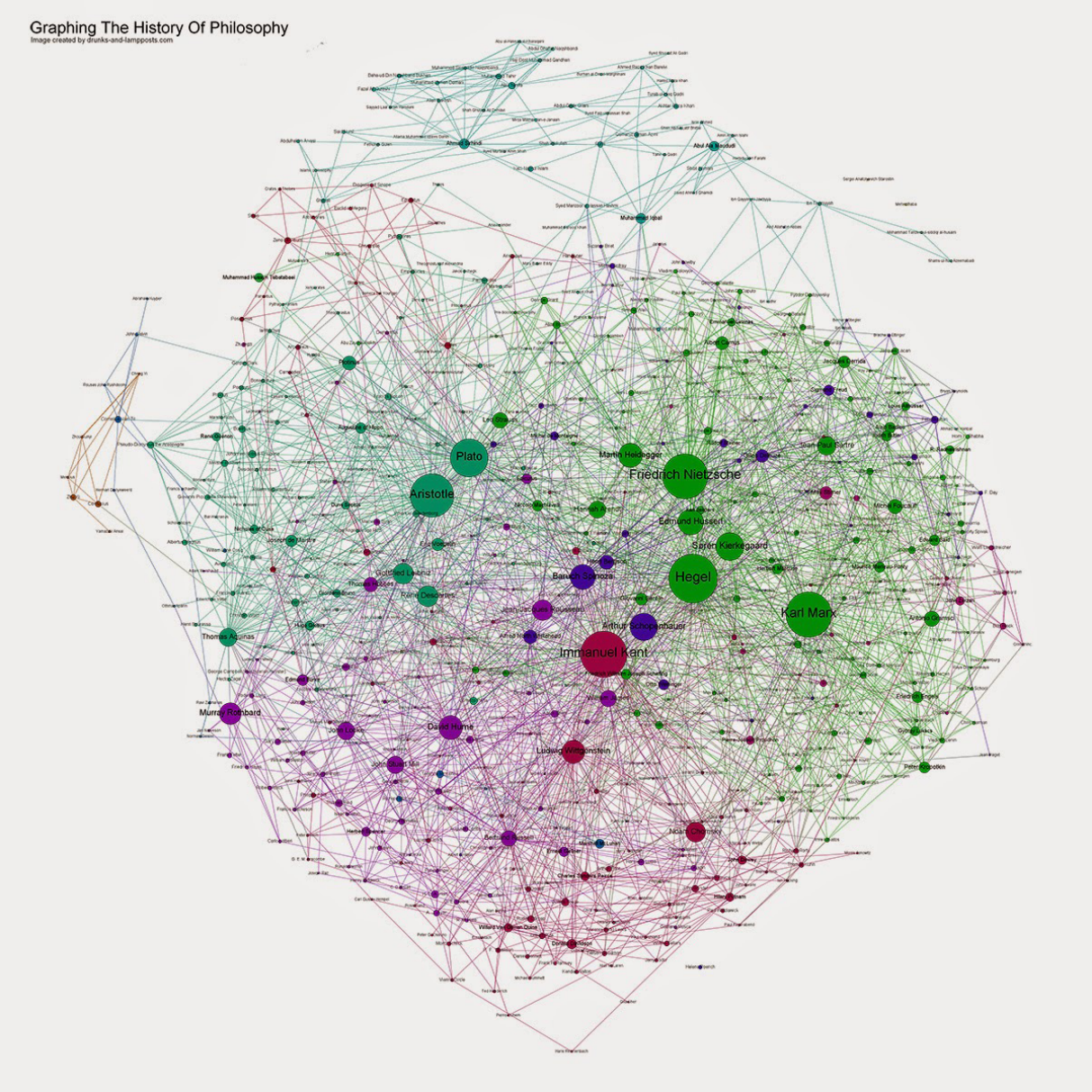


Figure A1.7. *Graphing the History of Philosophy*, by Drunks and Lampshots
Source: http://www.coppelia.io/2012/06/graphing-the-history-of-philosophy/

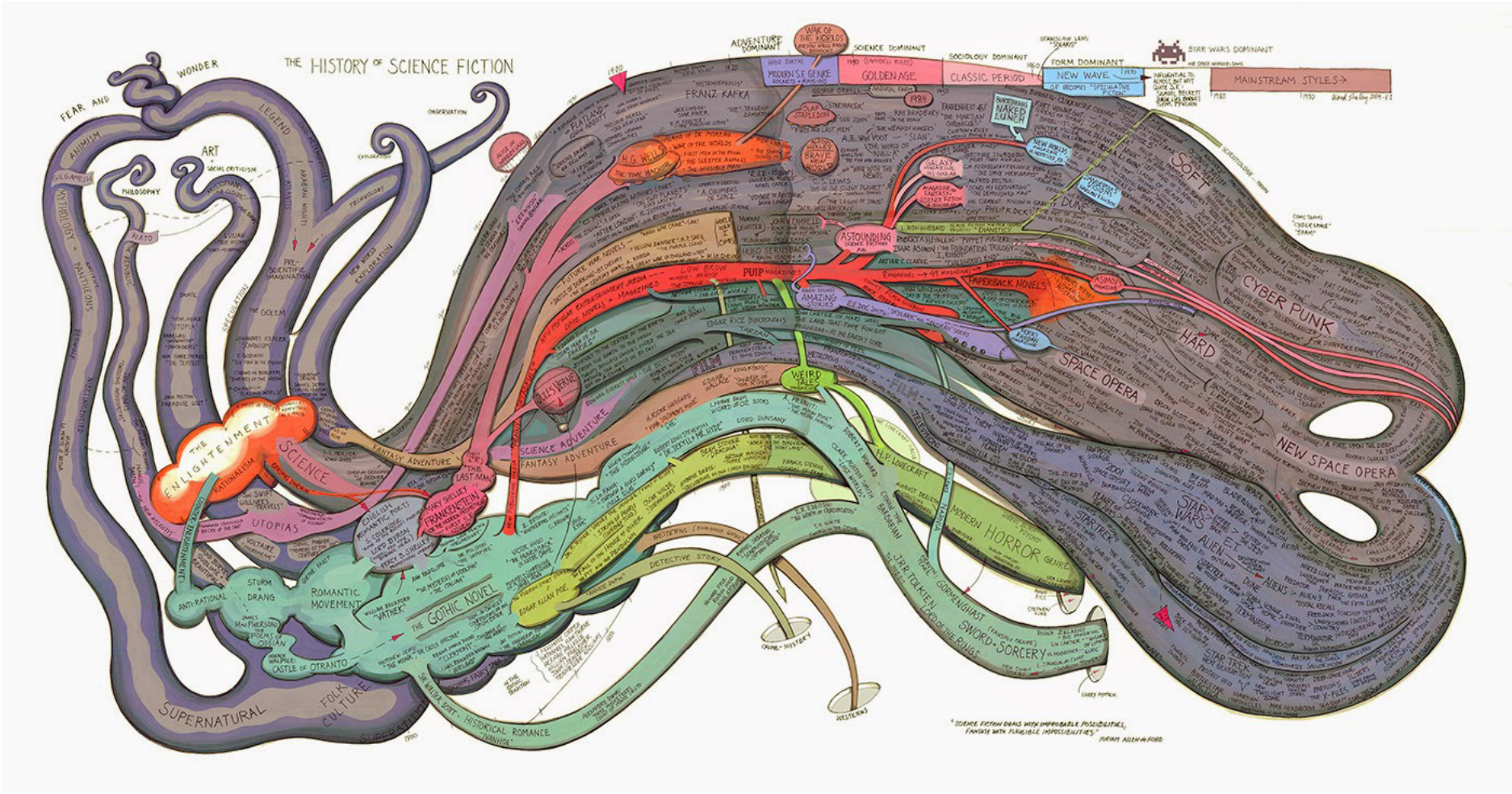


Figure A1.8: *The History of Science Fiction*, v.1, by Ward Shelley. Source: http://www.wardshelley.com/paintings/pages/HistoryofScienceFiction.html

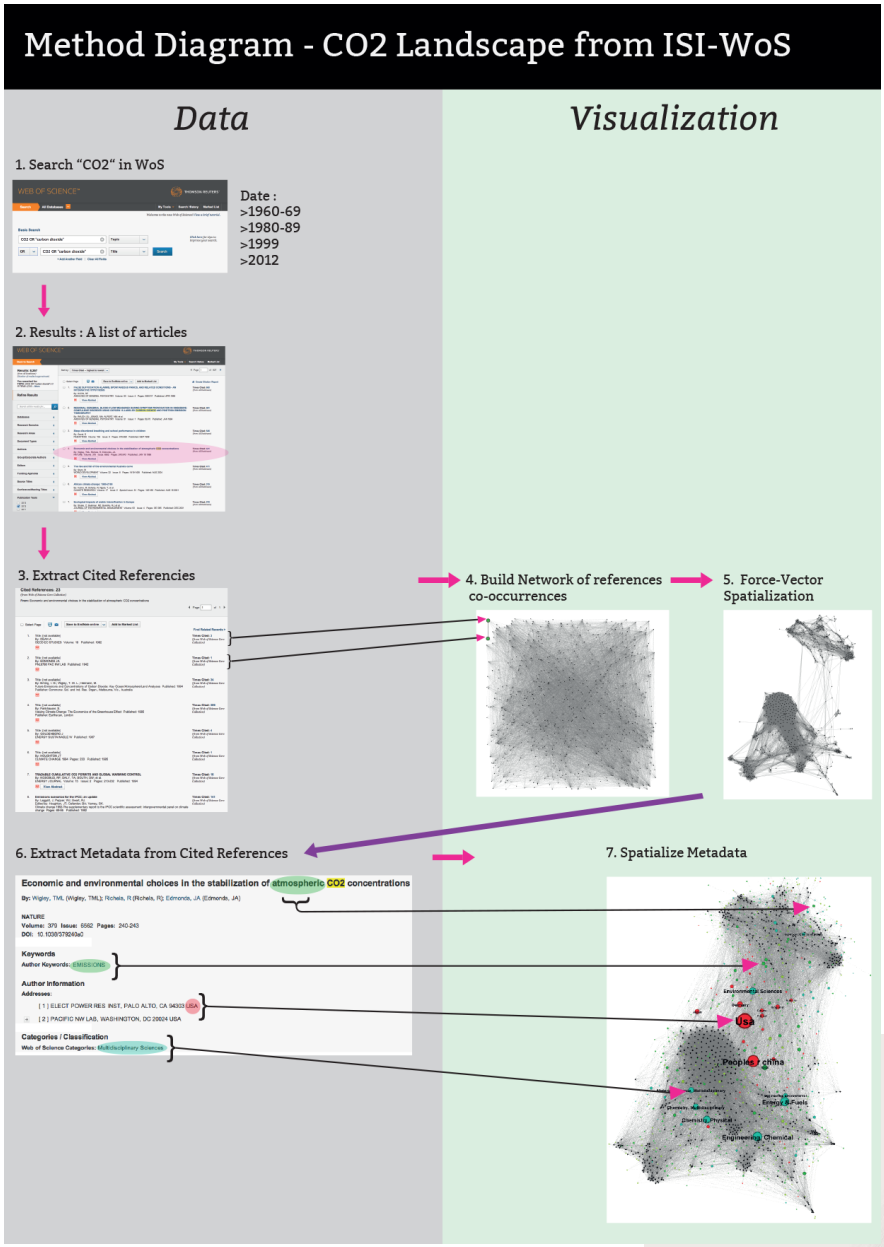


Figure A1.9. (top) *CO2 Landscape from ISI-WoS, Method Diagram*. By Kari De Pryck & Tommaso Venturini, forthcoming.

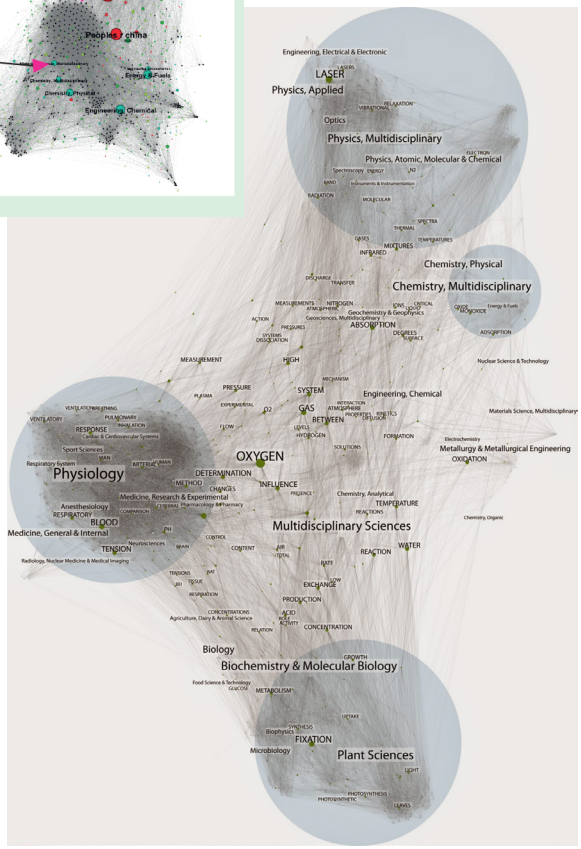


Figure A1.10. (right) *Carte Scientometrique des CO2*, by Kari De Pryck & Tommaso Venturini, forthcoming.

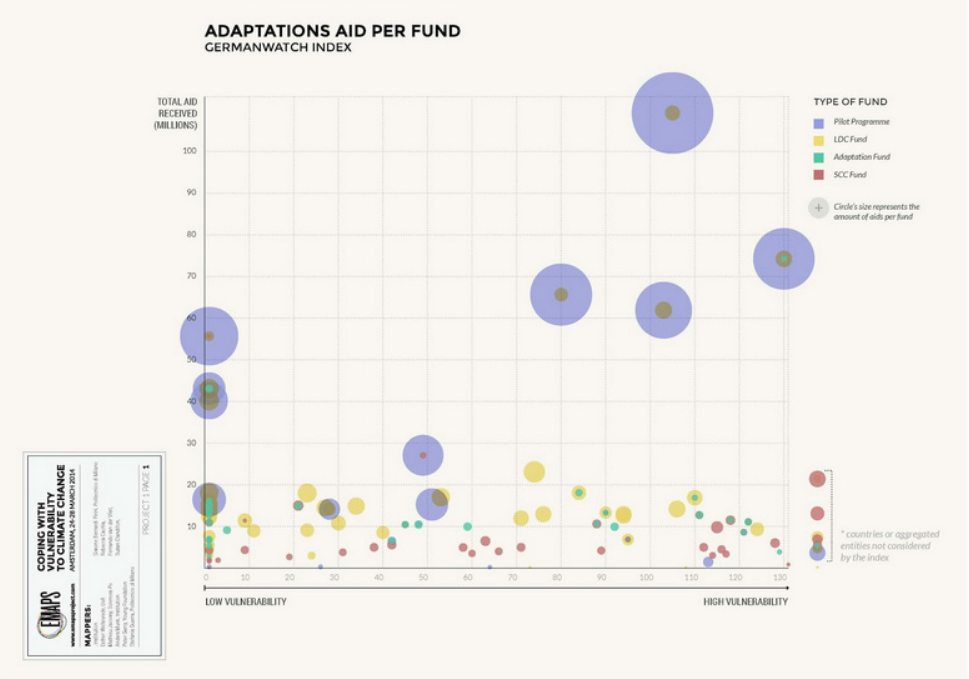


Figure A1.11. Adaptation Aid per Fund - Germanwatch Index. Source: Emaps archives.

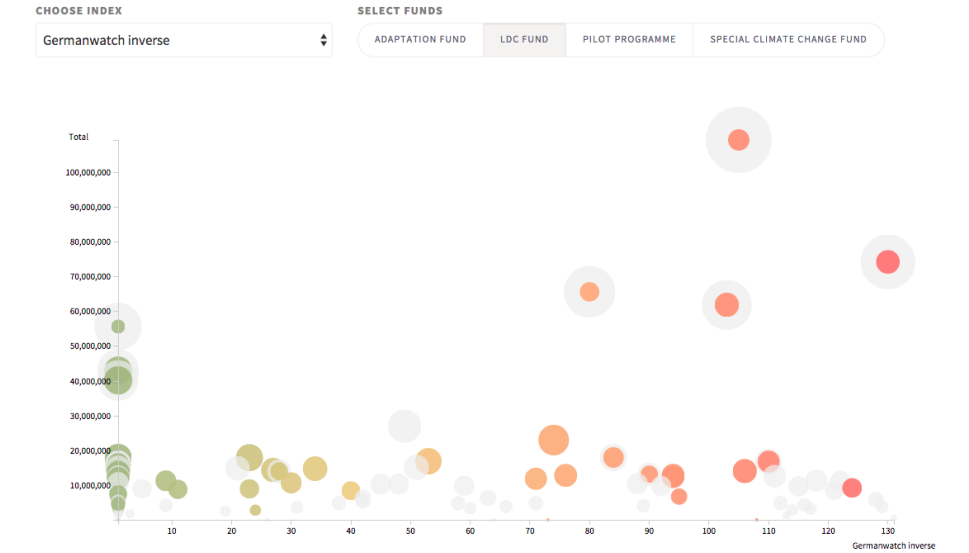


Figure A1.12. Multilateral Adaptation Funding And Vulnerability Indexes. Source: http://climaps.eu/#!/map/multilateral-adaptation-funding-and-vulnerability-indexes

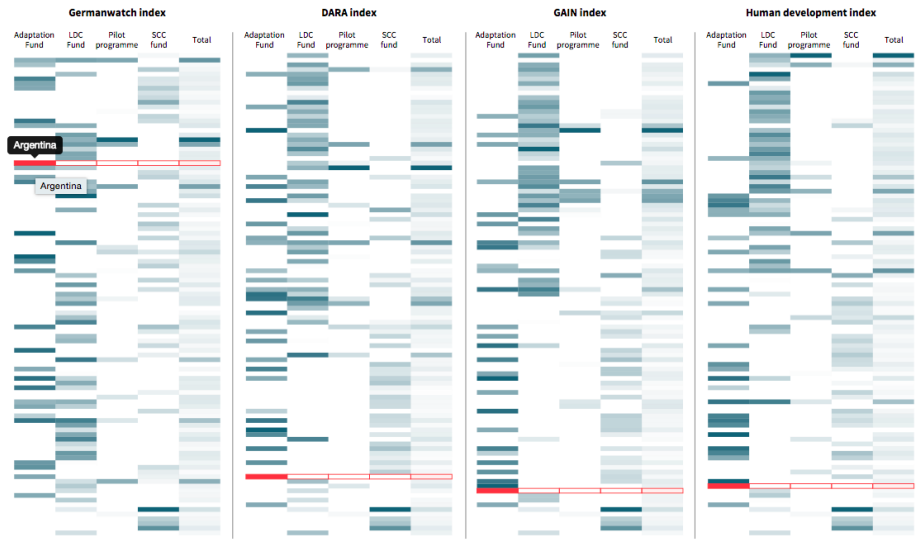


Figure A1.13. Multilateral Adaptation Funding And Vulnerability Indices - Matrix. Source: http://climaps.eu/#!/map/multilateral-adaptation-funding-and-vulnerability-indices-matrix

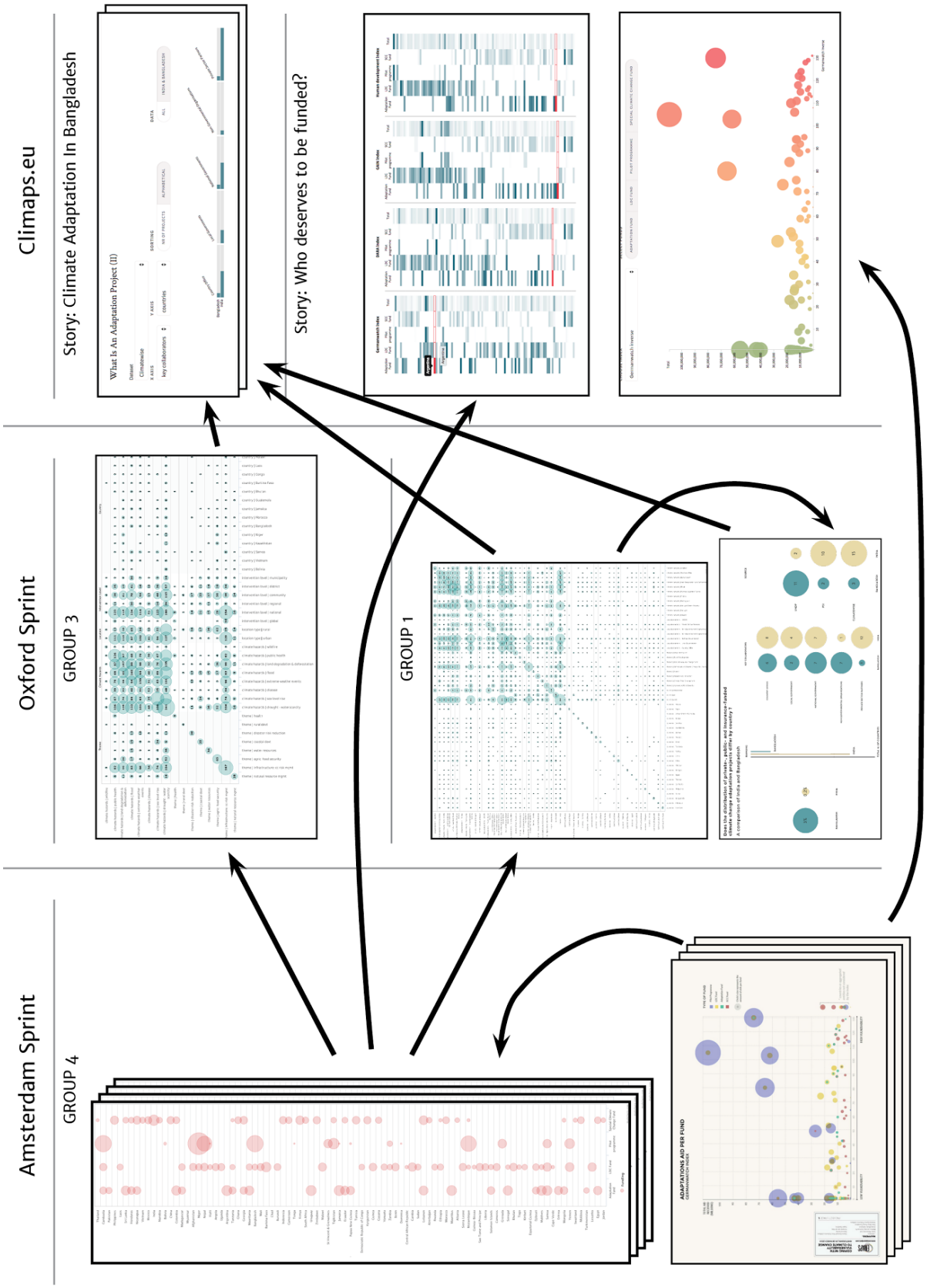
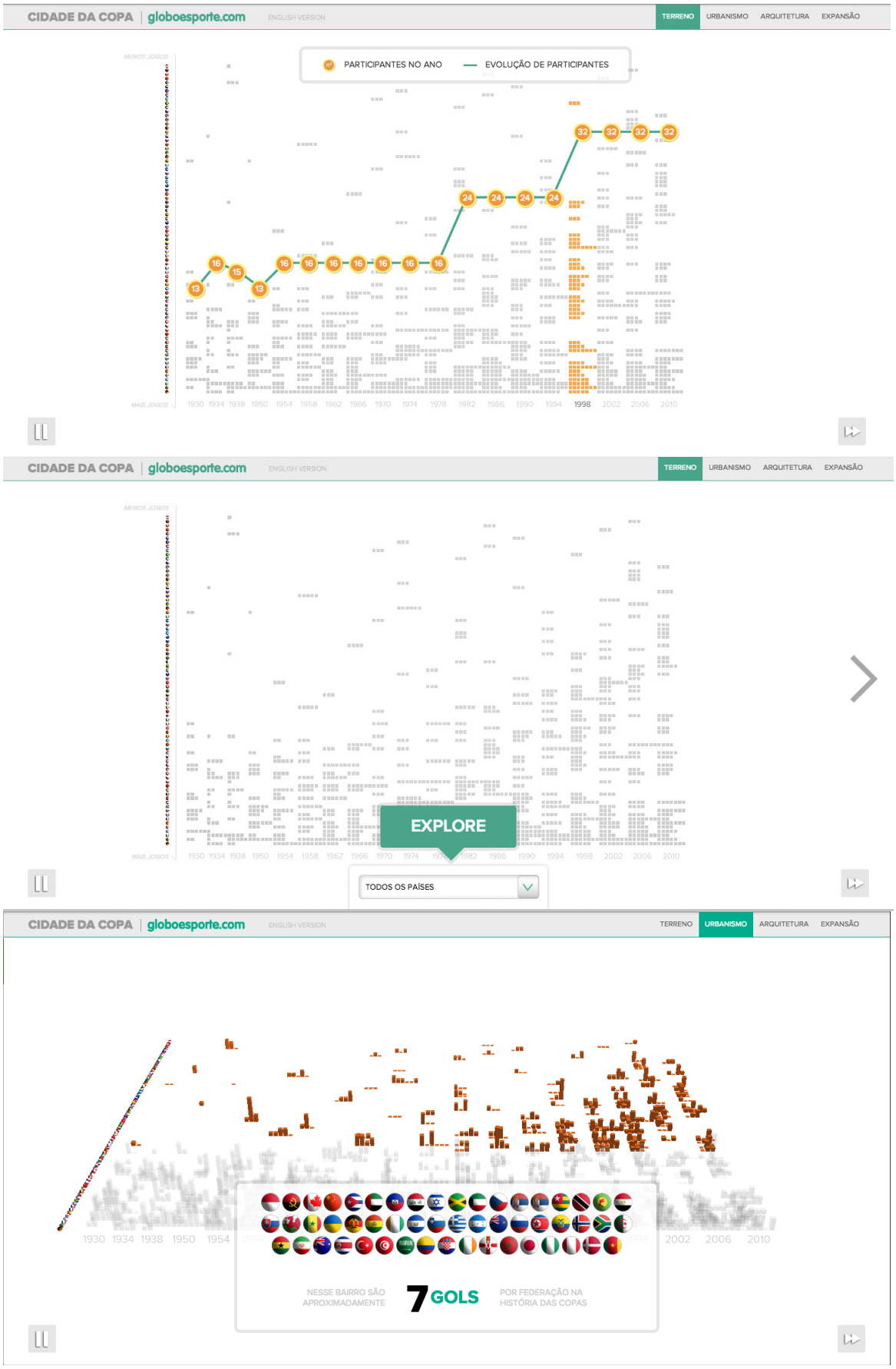
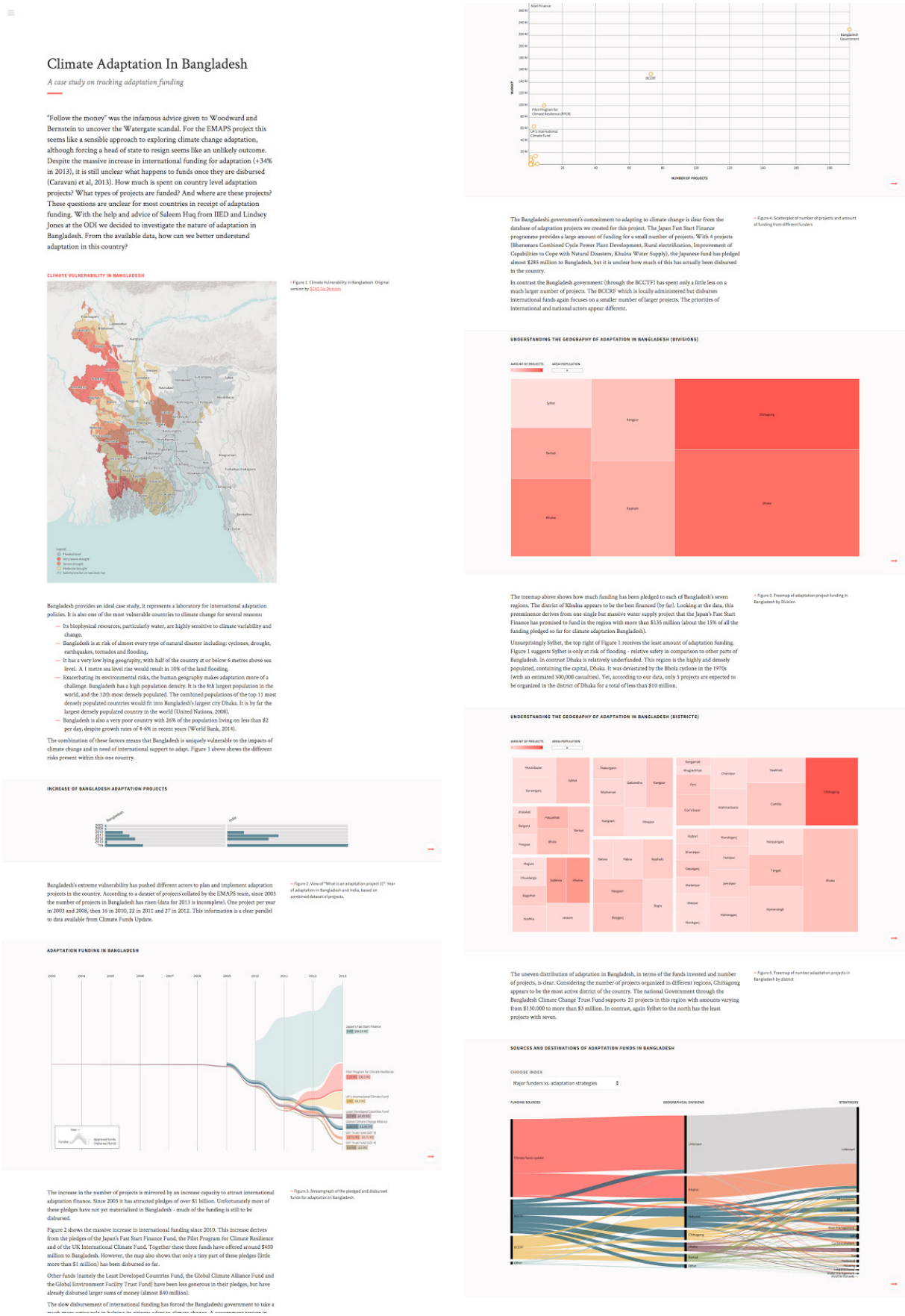


Figure A1.14. A schematic representation of some of the transformations observed in Emaps. Source: Emaps archives, Climaps.eu.



Figures A1.15, A1.16 and A1.17. *Cidade da Copa*. Source: http://app.globoesporte.globo.com/copa-do-mundo/cidade-da-copa/



Figures A1.18. *Climate Adaptation in Bangladesh*: Selected parts of the issue story. Source: http://climaps.eu/#!/narrative/climate-adaptation-in-bangladesh