Beyond immediacy and transparency.
A semiotic approach to discursive and rhetorical strategies in media visualization and data visualization

punctum.gr

BY: **Valentina Manchia**

ABSTRACT

Media visualization based on big cultural data, as "visualization without reduction" (Manovich 2010) is supposed to make data immediately and completely available, in contrast to classic data visualization, which visually translates information by means of "graphical primitives." On the other hand, from a pure functionalist point of view, also the visual form of diagrams, charts, and graphs, being fully proportional to the data values it conveys, is transparent with respect to its object (cf. Tufte 1990, 1997, 2001, and Card, Mackinlay, Shneiderman 1999).

In this paper we will try to consider both media visualization and data visualization (across several examples, including some Manovich's and Accurat's projects and *New York Times* graphics) as complex visual communication artifacts, not only from a purely informational point of view but from a semiotic point of view, by introducing a semiotic reflection on what we have proposed to call "discourse of data" (Manchia 2020a).

From our perspective, situated in the methodological framework of visual semiotics, and of the semiotics of scientific discourse, it might be interesting to pay attention to the whole process of constructing knowledge (and visual information) from data, understood as a chain of "devices of visualization" (Bastide 1985a, 1990 [1985b], 2001), investigating data as a channelled result, and also visualization strategies of specific–and oriented–discourses across data.

ARTICLE INFO:

Volume: **08**

Issue: **01**

Summer **2022**

ISSN: **2459-2943**

DOI: **10.18680/hss.2022.0006**

Pages: **85-113**

KEYWORDS:

Media visualization

Data visualization

Information design

Visualization

Translation

Punctum. International Journal of Semiotics | **08**:01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

In one of his more programmatically oriented essays, *What is visualization?* (2010), Manovich points out that media visualizations based on big cultural data, the huge media databases under his analysis, could more simply be called *direct visualization*. This is because visualizations of such large datasets would operate without mediation, presenting all visual data simultaneously and synoptically in a single visualization, a kind of image of images.

In a recent volume by the same author, which contains both *What is visualization?* and other cultural analytics essays, an effective portrait of his personal approach to visualizing big cultural data emerges. Such a portrait is set directly against information visualization, through the device of rhetorical questioning:

> Can we explore and study collections of cultural media and records of cultural behaviors *without system of categories* that languages impose on reality? Can we *avoid the quantification, measurements, and summarization* that comes with the use of statistics? Can we study big cultural data *without using numbers?* (Manovich 2020: 11)

The answer to these questions is precisely direct visualization, described here as pure access to cultural data without the interposition of any linguistic or discursive, statistical, or quantitative mediation. In direct visualization, data would not be narrated, measured, or schematized. Data would be made immediately and completely available, *as they are*. A clear example of this way of proceeding is illustrated in the essay on the *Visualizing Vertov* project:

> […] we don't measure or count anything. Instead, we arrange the sampled frames from a film in a single high-resolution visualization in particular layouts. This use of visualization without measurements, counting, or adding annotations is the crucial aspect of my lab's approach for working with media data sets […]. (Manovich 2013: 5)

In short, media (or direct) visualization seems to be a way of looking at images as a direct and immediate source of data. In media culture, data are the images themselves, and media visualization aims to make them accessible as they are to allow for subsequent operations, both quantitative and qualitative, on the totality of the images themselves. As in the case of the essay on *Visualizing Vertov* from which the quotation above is taken (and to which we will return later), media visualization generates "images of images" (Manchia 2014), which in turn can be the starting point for other images, for example, diagrams and derived pattern visualizations.

Dondero, who calls these types of visualization "mosaics" (Dondero 2016), later referred to "aggregates" of media visualization, from which "the observer can perform actions on the images, go back to the visualization's source documents, and so on" (Dondero 2017: 211), shifting from one mode of existence to another. In this perspective, the set of original images is virtualized when a pattern between and across images emerges.

If, then, media visualization, due to its complex nature as an "image of images," is a set of virtualities, from the point of view of an observer taking it as a starting point for subsequent investigation and interpretation paths, does it suffice to interpret it without presupposing any discursive mediation that structures it as a signifying object?

In other words, should such large datasets of images (such cultural Big Data) be considered only as a pure display of "a corpus as *artifacts*," with the sole aim of representing "the structure by showing *all* the elements which constitute it"? (Dondero 2020: 103). Would such collections of images merely constitute a virtual system (and a non-text) from which to actualize specific visual texts?

In fact, the adoption of this point of view seems to overshadow that even the media objects of media visualization, before any further elaboration, can be considered a discursive construction. Likewise, as we shall see, it is also possible to investigate the discursive dimension also with respect to classical information design and data visualization.

That is why, from this theoretical perspective, we would like to introduce a semiotic reflection on what we have called "discourse of data" (Manchia 2020a).

Regarding data and information visualizations, it is possible on the one hand to consider both the dataset construction procedures and the definition of relevant information paths as rhetorical in themselves, as oriented to a specific argumentative purpose; on the other hand, it is possible to analyze the visual and communicative strategies used to construct, both from an informational and an expressive point of view, an image that is just as effective.

From this point of view, taking into account Bruno Latour's reflections on the construction of scientific discourse, through a continuous dialogue with Greimas's narrative semiotics, as well as considering both the semiotic reflection on the role of images in scientific discourse in the essential work of Françoise Bastide and Paolo Fabbri, and the methodological framework of the Greimasian visual semiotics, we propose the investigation, from a semiotic perspective, of both big data collections and data (or direct) visualizations based on them. That is, to bring attention to data, information, and its restitution by paying attention to the practices and strategies of putting data into discourse. Even in the case of the most seemingly simple and immediate data visualization we are always faced, from a semiotic point of view, with the production of a meaning-effect that we can try to investigate more closely.

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

1. Rhetoric of immediacy and rhetoric of transparency in media visualization and data visualization

First, it may be useful to point out that the definition of direct visualization as "visualization without reduction" actually places it in contrast with data visualization. It in fact works by reduction, that is, by translating information into a simplified visual form:

> In my view, the practice of information visualization from its beginnings in the second part of the 18th century until today relied on two key principles. The first principle is reduction.
> Infovis uses graphical primitives such as points, strait lines, curves, and simple geometric shapes to stand in for objects and relations between them–regardless of whether these are people, their social relations, stock prices, income of nations, unemployment statistics, or anything else. […]
> Do all […] different visualization techniques have something in common besides reduction? They all use spatial variables (position, size, shape, and more recently curvature of lines and movement) to represent key differences in the data and reveal most important patterns and relations. This is the second (after reduction) core principle of infovis practice as it was practiced for 300 years–from the very first line graphs (1711), bar charts (1786) and pie charts (1801) to their ubiquity today in all graphing software such as Excel, Numbers, Google Docs, OpenOffice, etc. (Manovich 2010: 5-7)

Manovich summarizes in this way the fundamental coordinates for there to be, in his view, visualization, according to the meaning given to this term by classical information design. These are the use of spatial variables (position, size, shape) in the construction of a layout capable of best representing data and relationships between data, and the reduction to the minimum of the visual elements used (graphical primitives). Incidentally, a position not far from that of Jacques Bertin, author of *Sémiologie graphique* (1967), who almost forty years earlier had spoken of the relationship between content and expression, in information design, in terms of "transcription" by means of "visual variables."

On the contrary, media visualization approaches data directly, without any diagrammatic mediation:

> Rather than representing text, images, video or other media though new visual signs such as points or rectangles, media visualizations build new

> representations out of the original media. Images remain images; text remains text. […] In direct visualization, the data is reorganized into a new visual representation that preserves its original form. (Manovich 2010: 12)

As can be seen from these lines, diagrammatic mediation is radically opposed to the immediacy of direct visualization, both because it does not operate by reduction and because all data are immediately available. Access to data, in fact, is not mediated by any form of translation of logical relations by means of spatial relations (Peirce (1931-1935: 347 [4,348]) but occurs, we might say, by *ostension*: visual data are not returned through the interpretation of an expert, they are not synthesized by the position of points on a diagram but are simply made available "as they are."

Instead, the functionalist vulgate of data and information visualization turns out to be centered on a rhetoric of transparency, based on just the kind of graphic-visual mediation that the cultural analytics proposal seeks to overcome. In this perspective, a diagram, a schema, an infographic, a data visualization are invisible objects in themselves, in the same way that good typography is, for Beatrice Warde (1930), as transparent as a crystal goblet, which is even more perfect the better it allows one to appreciate every nuance of the wine it contains.

In short, for classical statistics and information design (cf. Tufte 1990, 1997, 2001, and Card, Mackinlay, Shneiderman 1999) the visual form of diagrams, charts, and graphs, being fully proportional to the data it conveys, and regulated by specific encoding rules, cannot but be transparent with respect to its object.

Emblematic in this regard is the so-called "data-ink ratio" codified by Edward Tufte, among the leading theorists of information design: "Every bit of ink on a graphic requires a reason" (Tufte 1983: 96; see also 96-105 and 123-137). According to this principle, the amount of ink used for data display should always be strictly proportional to the amount of information conveyed.

Given these premises, and the features we have pointed out (in direct visualization, the so-called display of the totality of the data itself; in data visualization, the access to pure data through graphical transcription), the visual transposition of data by direct visualization and data visualization nevertheless seems to look, in both cases, at the model of an immediate restitution of data.

More specifically, both the rhetoric of direct visualization and that of data visualization, following the two different directions we have mentioned, are hinged on the proposal of a discourse that is as objective and referential as possible, in which the marks of presence of the enunciating subject are declaredly erased (direct visualization), or minimized to the point where they can be ignored (data visualization).

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

2. From invisible to visible: visualization as a translation process

It seems interesting to us, from a semiotic perspective, to rather look at the restitution of data and information in its specificity as a "syncretic manifestation of a tracing activity" that combines texts, data, images, and diagrams (Fabbri 2001: 14, translation ours). The reference here is to Françoise Bastide, a biologist and a semiologist, close to Latour but a fellow student of Greimas too, and her work on scientific discourse, which narratively traces the mechanisms presiding over the construction of the scientific text, including its "iconography" (Bastide 1990 [1985b]), from a perspective that we also feel crucial for thinking about the "discourse of data" (Manchia 2020a).

Diagrams, schemata, and visualizations are not in fact a direct product of the data they refer to, at least not as much as functionalist information design would seem to suggest. On the contrary, following Bastide (1985a, 1990 [1985b], 2001), it is the device of visualization specifically developed by the enunciator that constructs, by making it "visible", its object.[1]

In particular, Bastide points to how in scientific texts very specific communicative strategies are at work. In more details, in scientific texts we are dealing with a particular type of enunciation that correlates an informative doing, which is always the dramatization of an action (the exposition of a laboratory result, the argumentation of a thesis) with a persuasive doing, where the enunciator presents himself as a witness to the real world and what happens in it, without speaking to assert his own identity (Bastide 1989 [1981]: 126-127).

In fact, every text, from a semiotic perspective, contains not only what it speaks about but, as if in a watermark, also its specific way of conveying and communicating its content–that is, it exhibits the traces of its production (of its enunciation). In this sense, any text, no matter how objective it pretends to be, can always be traced back to the communicative strategies put in place to construct this effect of immediacy or transparency.

Therefore, if the informative practice of the subject of scientific discourse is the construction of the referential level underlying every scientific discourse (see, in particular, Latour 1987, but also Latour, Woolgar 1986 [1979]), proper to its being both testimony and description of the so-called "real world," it is equally important to remember, at the same time, the persuasive function intertwined with this action, and therefore to study the implicit instances of enunciation responsible for the credibility of the restitution of such "real."

---

[1] In Bastide 1990 [1985b] "dispositif de visualization" become "technique of visualization." Here is translated as "device of visualization", following Bastide's Italian edition (Bastide 2001), as technical and discursive *dispositif* at the same time.

In Bastide's vision, however, visualization is not just a diagram or a series of diagrams, a microscope photograph, or a graph, nor is it just the visual output coming at the end of the scientific discourse production. Actually, the whole process of constructing the scientific fact is for Bastide the result of a narrative path that is to all intents and purposes a "technique of visualization" from an "invisible structure" to a "visible structure" (Bastide 1990 [1985b]: 190).

In more detail, the assignment of a visible structure (e.g. a graph accompanying an article on a new scientific discovery) to an invisible structure (the object of the scientific discovery as such) is only the final step of a much more complex process of *translation* (Bastide 1985a), in connection with the concept of *inscription* in science (Latour, Woolgar 1986 [1979]).

According to Bastide, such a translation process starts with the channeling of the scientific observation in a given direction, by identifying against the background of the so-called "nature" a scientific object detected and measured in the laboratory by an "Operating Subject", that is the researcher with his/her technical devices and filters (Bastide 1990 [1985b]: 190). This working object, which is therefore first and foremost an oriented collection of laboratory data, must then be translated again into a specific discourse in order to become an official "scientific object" within its scientific community of reference (Bastide 1985a).

In other words, between the so-called "natural element" and the final "object of knowledge" we need to hypothesize a dual process of visualization based on two different–but related–devices, i.e. the experimental conditions (laboratory technologies, working hypotheses, but also the application of categories of analysis and description) that enable the researcher to "see" his/her object, and the textual, verbal, visual and diagrammatic devices that enable the researcher to "make visible" his/her experimental results to the reader of the article.

> For convenience, one can conceptualize the article as containing two distinct layers of text: the description of the experimental technique that enabled the results to be obtained, and the general framework of the problem of the structure of transfer RNA.[2] As long as it is not contested the two layers of the article form a whole that I will call a technique of visualization. I would like to stress one of the very general characteristics of the article; it is a matter of using experiment to make one see what is invisible. The whole article is a technique of visualization for the public, for the scientific community concerned; the experimental conditions concern a technique of visualization more limited to the researcher' s own usage. (Bastide 1990 [1985b]: 188-189)

[2] "To introduce my argument, I will take as an example a recent discussion in Nature concerning an accusation of fraud that was brought against an article thirteen years after the original publication. The article under allack describes the crystallization of a transfer RNA for valine extracted from yeast" (Bastide 1990 [1985b]: 187).

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

In this final stage of translation, images–photographs, tables, graphs, visualizations–play a key role, and it may be worthwhile to attempt an "iconography" of them (Bastide 1990 [1985b]), within the methodological framework of Greimasian visual semiotics (Greimas 1984; Floch 1985).

In our vision, Bastide's position is thus doubly interesting when transposed from the scientific discourse to the "discourse of data." On the one hand, it allows us to investigate how the so-called data become data, that is, through what procedures researchers construct datasets as starting points for their work, somehow narratively governing the standing out, in the discourse, of such information, through the affirmation of given categories of analysis to the detriment of others.

By quoting Johanna Drucker's work on visualization and interpretation (2020), data can rather be considered as *capta*, that is, not as "preexisting entities" (as a fact of "nature", Bastide would say), but as the captured result, that is "the products of a process of parameterization and modeling."[3]

On the other hand, Bastide's position invites us to analyze visualization (both media and data visualizations) not as an immediate presentation or a transparent transcription of data, but as a visual translation of a specific–and oriented–discourse across data.

In the next sections we will try to consider both media and data visualization as complex visual communication artifacts, not only from a purely informational point of view, but also from a semiotic point of view. In other words, we will try to reconstruct visualization devices and related discursive strategies related to both content organization and data display.

From a methodological point of view, the semiotic analyses in the next paragraphs will be at the intersection of narrative semiotics dealing with complex syncretic texts, including scientific and diagrammatic ones, and visual semiotics. Specifically, some examples of direct visualization, coordinated by Lev Manovich (*Visualizing Vertov* and *On Broadway*, section 3) and data visualization, from Accurat's *The Room of Change* to *New York Times* visualizations of Covid-19 (section 4), will be analyzed.[4]

[3] About design as translation activity see also Zingale (2016). Starting from a "problematic objectuality", from which a design process starts (Zingale, 2012), the author suggests a translation model that connects a "briefing-text", with content and data analysed and textualized, to an "artefact-text", that is "content and data translated into an artefact" (Zingale 2016: 9-10).

[4] Regarding scientific and diagrammatic images from a semiotic perspective, see in particular Fabbri (1998a), (1998b), (2001), (2014), (2021), Stjernfelt (2007), and Dondero and Fontanille (2012). For an argumented exploration of the field of visual semiotics see Mengoni (2021).

3. Images remain images? Discursive strategies in some examples of direct visualization

A representative case is the already mentioned *Visualizing Vertov* project, carried out by Lev Manovich and his research unit at the time (Software Studies Initiative, at New York University) focusing on the analysis of the work of Tziga Vertov. Among the methods used, alongside digital image processing software capable of measuring the visual properties of each frame of Vertov's films, was precisely direct visualization.

The goal of the project, published in 2013 in digital form on the research group's website, was to investigate, from the perspective of what is programmatically called *exploratory visualization*, the formal structures underlying Dziga Vertov's production.

Manovich decides to focus on the author known for his theory of *Kino-Eye*, of cinema as overcoming the human limits of vision. *Visualizing Vertov* is in fact an attempt to look inside the *Kino-Eye* but in reverse, peering into the workings of the gears that preside over filmic construction. This is possible, in particular, through visualizations of visual data extrapolated from frames and sequences that explore, in unprecedented ways, certain visual features of the films. This is the case with the image that collects, in a single glance, all the sequences of *The Eleventh Year*, represented by the second frame of each, and that shows the alternation in both the editing and the formal components of the sequences, such as the progressive rotation of face shots (fig. 1). Another

Figure 1. Lev Manovich, Software Studies, *Visualizing Vertov*, media visualization of each of 654 shots in *The Eleventh Year* (Dziga Vertov, 1928).

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

Figure 2. Lev Manovich, Software Studies, *Visualizing Vertov*, visualizations of movement in three shots from *Man with a Movie Camera* using frame averaging techniques.

visualization condenses several frames into a single frame to visually represent the amount of movement present in a sequence of *Man with a Movie Camera* (fig. 2).

A deeper exploration of *The Eleventh Year and Man with a Movie Camera* shows recurring patterns in the relationship between some visual qualities (light, amount of movement, change or stillness of elements in a sequence) and other parameters that derive from the director's stylistic choices (the duration of a scene, the structure of the montage). From the "mosaics" in media visualization thus derive diagrammatic visualizations such as the one in Figure 3, in which the bar chart seems to reflect the principles of gradualism and radical opposition dear to the montage theories of the time, through the rhythmic and abrupt visual changes between sequences.

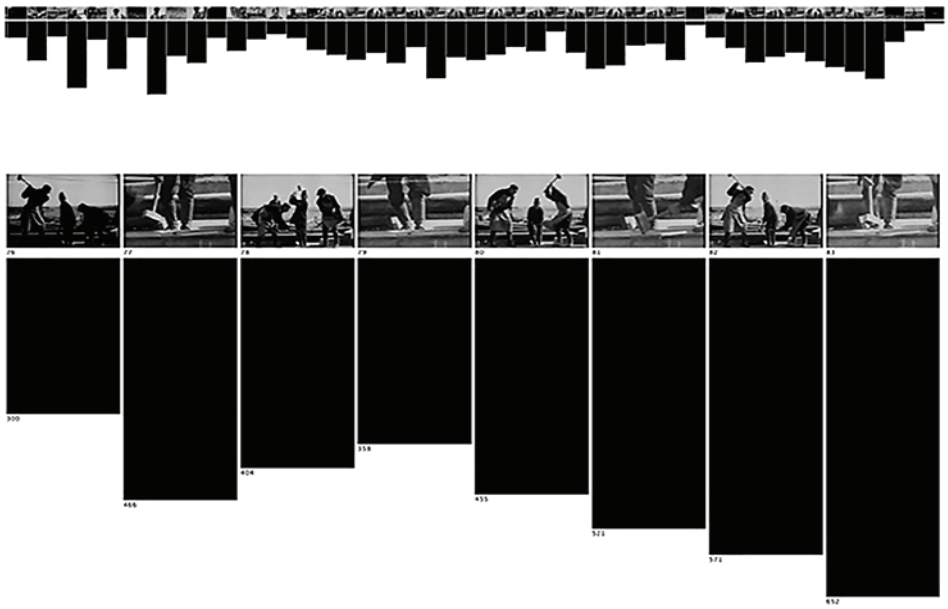

Figure 3. Lev Manovich, Software Studies, *Visualizing Vertov*, average amount of visual change in every shot in *The Eleventh Year*. Each bar represents one shot. The length of a bar corresponds to the average amount of visual change in the shot.

An initial observation concerns the ways in which these and no other direct visualizations were constructed. In fact, as Manovich (2020) himself pointed out, direct visualization is not only a way to show all images in a collection in a single visualization. It is also the result of the skillful and organized management of the metadata describing the images and the information they carry–and it is always possible for researchers to add additional layers of analysis that can make the visual corpus seen in a specific way, leading to different direct visualization outputs. In fact, "media visualization offers a new way to work with such information," where this information resides in the specific way the researcher looks at the images.[5]

A second observation is the quite peculiar argumentative strategy with which Manovich conducts his Vertovian investigation, reducing the text to a mere commentary on the visualizations, thus reversing the traditional structure of the academic paper with images accompanying the text.

Third, the visual argumentation around Vertov lines up very different visualizations, from an initial general chart, showing the incidence of montage in 20th-century Russian film production, to visualizations that progressively go into detail about the director's work, constituting a series with a "progressive focusing of attention" (Bastide 1990 [1985b]: 196) on the main details. The aim is for visualizations to first speak "from a 'bird's eye' view of the cultural artifacts (hundreds of 20th century films) and to gradually move closer and closer–similar to how Google Earth allows you to start with the Earth view and then zoom in and eventually enter a street view" (Manovich 2013: 5). Derivative data visualizations serve to stylistically compare the two films considered more closely, within the corpus, by showing the difference, for example, in terms of the length of sequences and their distribution, between *The Eleventh Year* and *Man with a Movie Camera*, made up of short and not-so-short sequences, appropriately adjusted and alternated.

The reader is thus faced with the construction of an oriented path between media visualizations and data visualizations, derived from an earlier in-depth study of Vertovian themes, to highlight an underlying discursive structure. In other words, it is from the identification of the elements to be compared, the films *The Eleventh Year* and *Man with a Movie Camera*, as two poles of fundamental importance in Vertov's work that the patterns the systematization of data helps to highlight gain prominence.

Considering what has been observed so far, the patterns that also emerge among the images are not just pure visual saliencies, but the result of skillful work

[5] "As I already explained, media visualization exploits the presence of at least minimal metadata in a media collection, so it does not require the addition of new metadata about the individual media items. However, if we decide to add such metadata—for example, content tags we create via manual content analysis, labels for groups of similar images generated via automatic cluster analysis, automatically detected semantic concepts (such as objects, types of scenes, or photographic techniques used), face detection data, or visual features extracted with digital image processing—all this information can also be used in visualization" (Manovich 2020: 224).

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

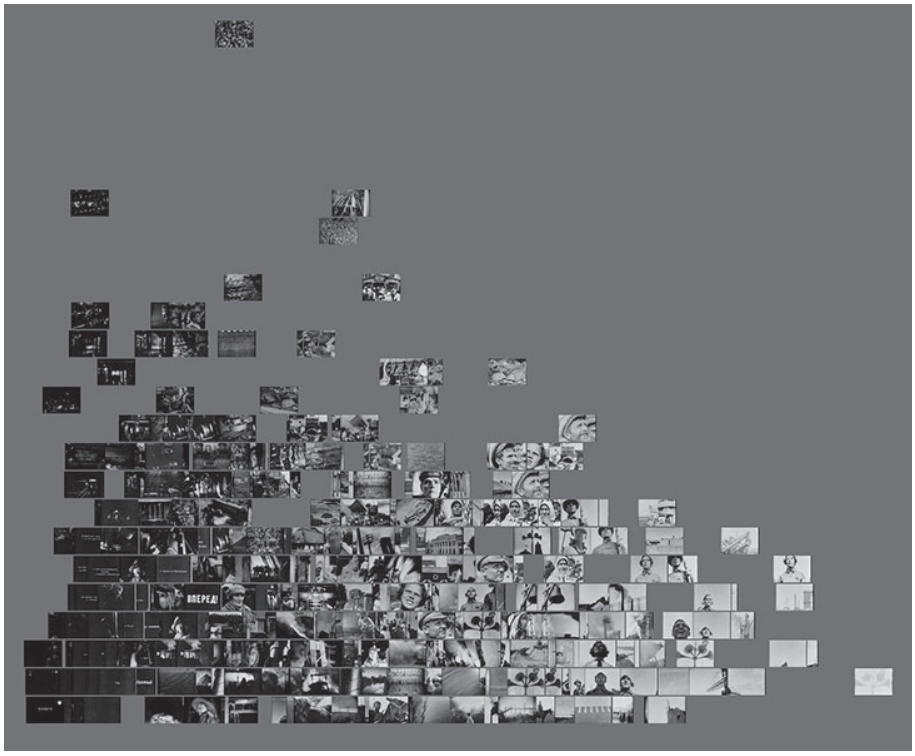

Figure 4. Lev Manovich, Software Studies, *Visualizing Vertov,* visualization of *The Eleventh Year* showing second frame of every shot. The frames are sorted their visual properties. X-axis–mean (average) gray scale value of a frame. Y-axis–number of shapes in a frame.

in structuring the source material (Tziga Vertov's films) according to specific directions of inquiry. A good example of this is *The Eleventh Year* visualization, with frames positioned on two dimensions according to their visual characteristics, their average gray scale values, and the number of shapes (Fig. 4). Manovich himself summarizes as follows the close relationship between media visualization and metadata and/or different combinations visual characteristics to use, "each creating a different 'map' of a film" (Manovich 2013: 26), even in a case like this where the source material, as it is for Vertov's corpus, is eminently visual:

> The portfolio of visualizations with commentaries below starts with three visualizations that plot the already familiar quantitative measure of films, i.e., shot lengths (1-3). It then moves to illustrate other possible strategies (4-9) using "media visualization" approach. To create media visualizations, sampled frames from a film are arranged by the software in different layouts using existing metadata (i.e., frame numbers, shot boundaries, manual annotations of shot content, etc.) and/or automatically created measurements of visual properties of images, shots and whole films. I don't include any numerical

> measurements of the films in my main analysis, in order to emphasize that visualization techniques can help to us explore films and identify patterns beyond mere statistics. (Manovich 2013: 15).

If media visualization can reveal patterns in the data, we could also say, at the same time, following Drucker (2020) and her approach to "modeling interpretation" that media visualization preliminary operations (selection of categories to be considered for analysis, of quantitative and qualitative dimensions to explore across the data) inform patterns.[6]

A second example of direct visualization is the interactive installation created for the New York Public Library by Manovich and his research group specifically as part of the exhibition "Public Eye: 175 Years of Sharing Photography" (December 2014-January 2016).

From the very beginning *On Broadway* was conceived in parallel with a web application (that is still available). The project as a whole was designed to map a specific portion of Manhattan, Broadway, in its entirety and over a period of time from 2009 to 2013, telling the story of both the area and the continuous and constant interaction of subjects (not only inhabitants, but also tourists or passers-by) with it (fig. 5).

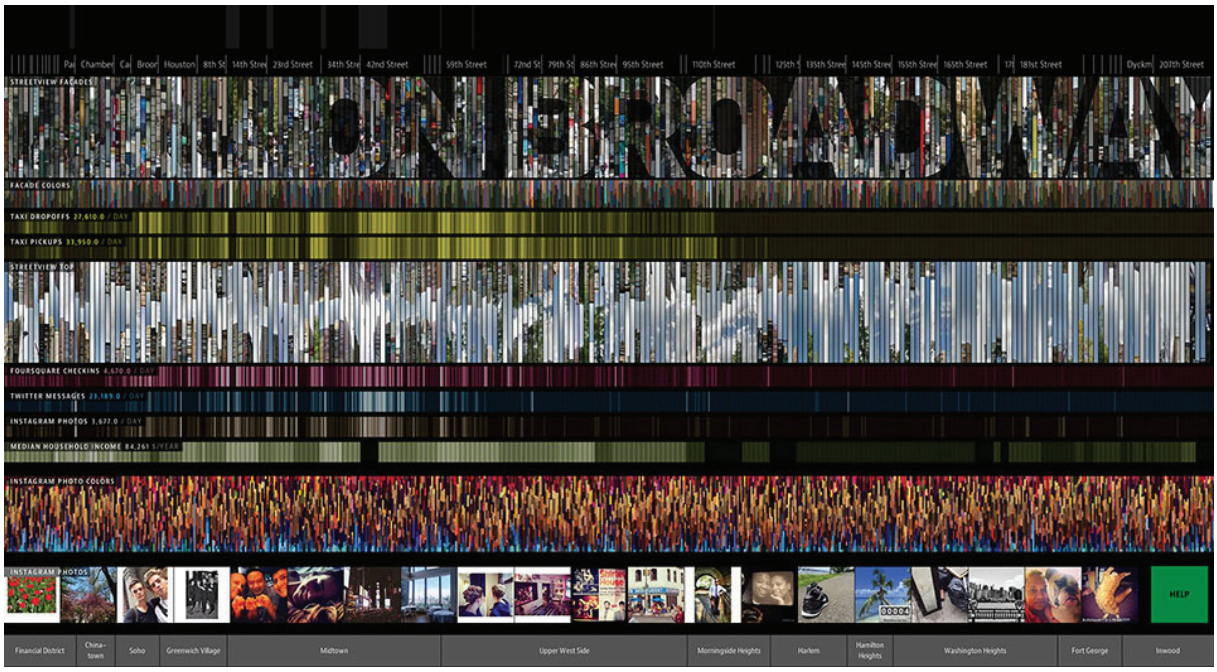

Figure 5. Daniel Goddemeyer, Moritz Stefaner, Dominikus Baur, and Lev Manovich, *On Broadway*, the view of the complete 13 miles of Broadway in Manhattan.

[6] About data, and classification, see also Crawford (2021).

Punctum. International Journal of Semiotics | **08**:01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

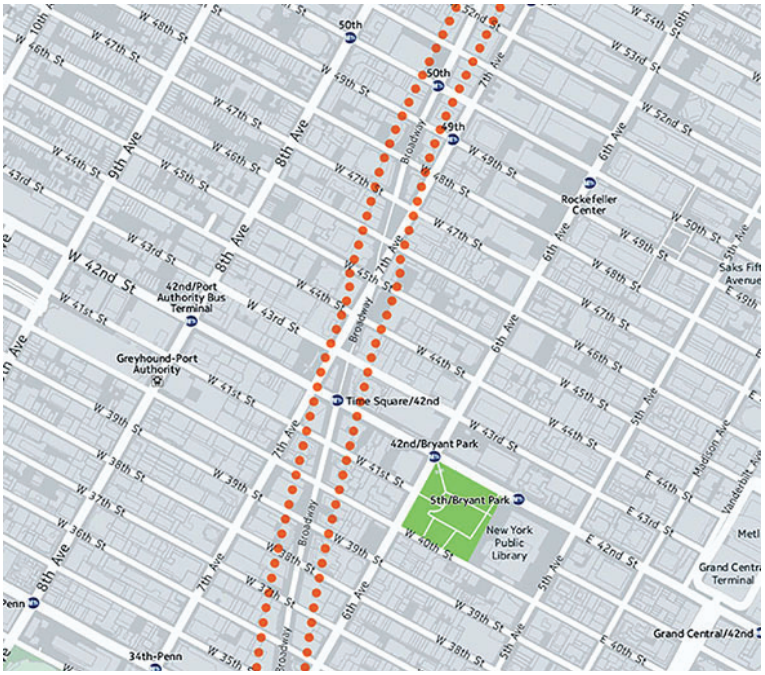

Figure 6. Daniel Goddemeyer, Moritz Stefaner, Dominikus Baur, and Lev Manovich, *On Broadway*, close-up showing the width of the area centered on Broadway used as data filter.

The project coordinates millions of data and thousands of images. Among the data fed into the project are photos shared on Instagram, posts with images from Twitter, images from Google Street View, Foursquare check-ins, cab ride data, and others. Data was collected extensively along Broadway, on both the left and right sides, and then channelled into multiple layers (the horizontal layers), each of which identifies a distinct survey category (Fig. 6).

Not all information is transposed in its entirety, via direct visualization: that is, of some datasets only a schematic graphical representation is provided, through diagrams or graphs of more or less classical layout, according to the rules of the most classic data visualization. Fully displayed, however, and in the true spirit of media visualization, are the images (visual data) collected in the field–and this is what we will focus on here.

We have already explored this example and in particular its restitution of data-images in the form of direct visualization, with a particular focus on the ways in which subjectivity effects are produced in texts such as these, informational and expressive at the same time, in which a complex articulation of semiotic forms coexists (Manchia 2020b).

Specifically, through the prism of Marin's (1983) reflection on the two possible polarities in cartographic representation, the modality of *narrative* and the modality of *description*, we emphasized how even in the immediate display of the images of *On Broadway* can be found the traces of an enunciational device at work in the visual levels of restitution of individual data.

In fact, the inclusion of the observer within the perspective from which the images originate is achieved not only through the interaction provided by the interface (the choice to actualize a single path of the many virtually possible), but also through the constant adoption of a bottom-up perspective in the so-called Streetview Top layer (the exact opposite of a zenith view, from above), which runs parallel to the Streetview Facades layer, the frontal images of the buildings along the street (in Fig. 7, see the first and second rows of photographic images, respectively).

In more detail, the first linear sequence of images (directly derived from the land point taken into consideration) of the facades of the Broadway buildings, in the Streetview Facades layer, has the effect of reproducing the continuity of the street facades in a way that recalls architectural elevation. The second linear sequence in the Streetview Top level, on the other hand, which runs parallel to the first, also welds to the Broadway elevation a kind of phantasmal skyline, since the images from the bottom upwards identify the height of the buildings as well, point by point, reconstructing for the user, by means of montage between the two levels, the overall experience of walking down Broadway.

While "images remain images" as they are collected at their respective points, according to the principles of media visualization, offering a virtuality of possible paths for the user to actualize, it seems more interesting to investigate how this set of images is constituted no longer as a set of pure virtualities but as a true signifying entity, even if not graspable in its totality.

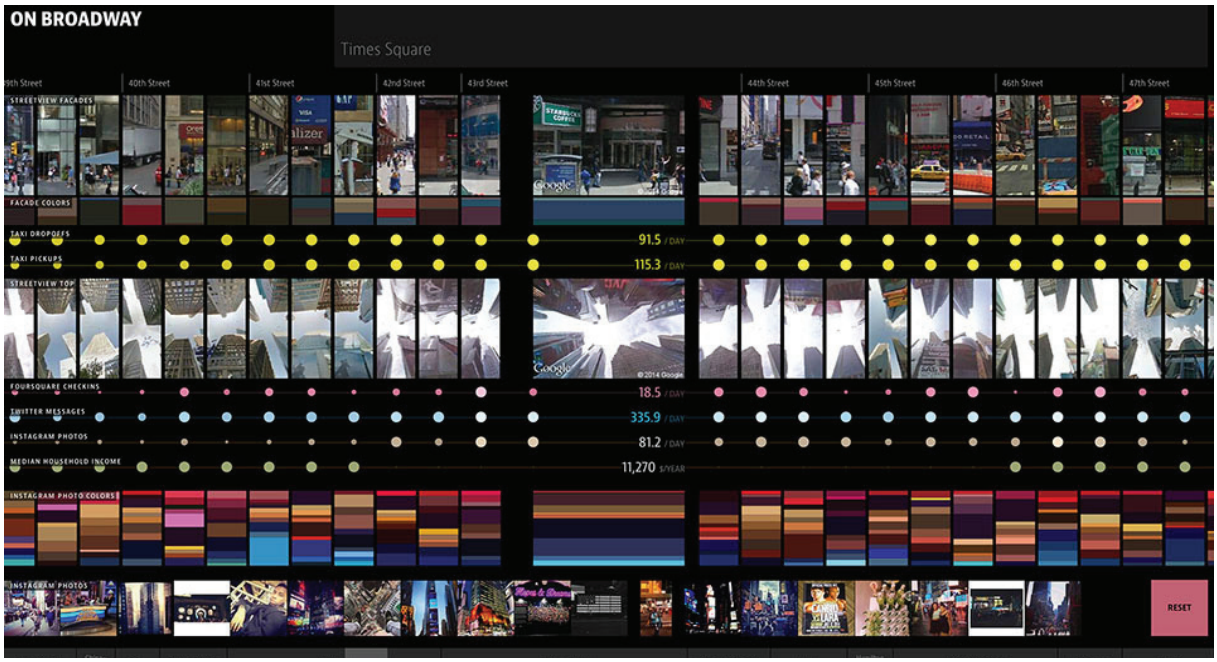

Figure 7. Daniel Goddemeyer, Moritz Stefaner, Dominikus Baur, and Lev Manovich, *On Broadway*, detail (Streetview Facades and Streetview Top).

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

If we consider it as the result of a series of selection filters and categorical procedures, thus as a cultural and semiotic construct, a collection of big data is in fact already in itself a real object, a text in its own right, precisely because it could not exist, even in the open form that it has, without the "reading grid" that has in fact structured it as such.

In the specific case of *On Broadway*, it is thanks to the capillary organization of data collection on the ground and on points at a short fixed distance from each other (every 25 meters or so)–a distance capable of measuring most effectively, according to a visitor's perspective, the architectural and thematic change along the two sides of the street–that the individual data will be able to offer such an almost "immersive" experience to the viewer. Likewise, in *Visualizing Vertov*, the multiple paths for the user to actualize from the virtualities offered by the source database would not be possible if they were not based on the choices necessary to create the database itself.

Returning once again to Françoise Bastide and her seminal investigation of devices of visualization, the entire database, insofar as it is structured and organized, is itself, too, the result of a specific "making-to-see", which is framed within the peculiar research perspective of the data scientist or researcher and takes its point of departure from it.

A similar critical position toward the pure virtuality of datasets is now increasingly evident, and in various fields of research, for example among policy and data scientists. To give a few examples, scholars currently working on the topic of data gaps point out how the so-called "raw data" are in fact the result of selection filters and structuring practices. In particular, some have shown the often-undetected incidence of big data's periphery (Lerman 2013) or of the so-called "dark data" (Hand 2020), through which existing phenomena remain effectively "invisible" because data cannot (or will not) be collected on them. Lerman, in this regard, speaks of the "nonrandom, systemic omission of people who live on big data's margins," and thus of "big data's exclusions" (Lerman 2013: p. 57). Others (Giest and Samuel 2020) proposed a categorization of data gaps based on different data availability for policymaking, from *primary data gap* (no availability) to *secondary data gap* (limited availability) to *hidden data gap* (no awareness of such availability). In this regard, some have also highlighted the bias implicit in the very existence of the big data phenomenon as such, as a technological-social phenomenon involving complex construction dynamics that are often overlooked (Boyd and Crawford 2012; Olteanu et al 2019; Crawford 2021).[7]

It is for this reason that, in our view, a semiotic approach to different data visualization and translation devices could enable us to better investigate and explore these kinds of dynamics and construction practices.

[7] For a semiotic overview on data gaps and data bias see also Manchia (2021).

4. Data-tapestries and data-walls. Discursive strategies in some examples of data visualization

The third example we will consider is instead ascribable to the other paradigm, that of data visualization and the rhetoric of transparency. It is *The Room of Change*, a "data-driven wallpaper" designed by Accurat studio specifically for the XXII Milan Triennale (March 1-September 1, 2019), *Broken Nature: Design Takes on Human Survival*, directed by Paola Antonelli.

The project, arranged in an entire museum room, is a complex installation, combining several artifacts (fig. 8),[8] amongst others two large screens projecting images taken from NASA's *Images of Change* series, which by their very position trace an imaginary corridor that directs the visitor's movements and attention toward the back of the hall.

At the back, on the three walls of the room, runs a large data visualization divided into three moments, roughly corresponding to the three walls: the left side wall, from

Figure 8. Accurat, *The Room of Change*, installation entrance (© La Triennale di Milano; photo Gianluca Di Ioia).

[8] A first analysis of this project, more focused on some enunciational issues, is in Manchia (2020b).

Punctum. International Journal of Semiotics | **08**:01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

100 A.D. to 1900 A.D.; the middle wall, from 1900 to 2000–and overflowing with the years 2000-2019 on the right wall, which accommodates the time frame from 2020 to 2400. Along each wall, from 100 AD to 2400, run eight themes (from 1. Nature to 8. Technology), eight parallel macronarratives, each characterized by precise patterns of lines and shapes.

Each macroseries of data is in turn structured into three subseries, each of which declines the central theme in its own way. The main themes are, specifically: 1. Nature; 2. Universe; 3. Animal Kingdom; 4. Society; 5. Hope; 6. Happiness; 7. Science; 8. Technology.

Each macroseries is then divided into three more specific subseries (for example, 1. Nature includes the three subtopics 1.1. The disappearance of the Aral Sea; 1.2. Human impact on the environment, 1.3. The effects of climate change), for a total of 24 subseries arranged in horizontal rows and running, linearly, along all three walls (fig. 9).

The visual modulations present on the expression plane, as might have been expected, are not developed in parallel with the themes of the macroseries, but are exploited in a very fine and always varying way to represent the development of each individual dataset.

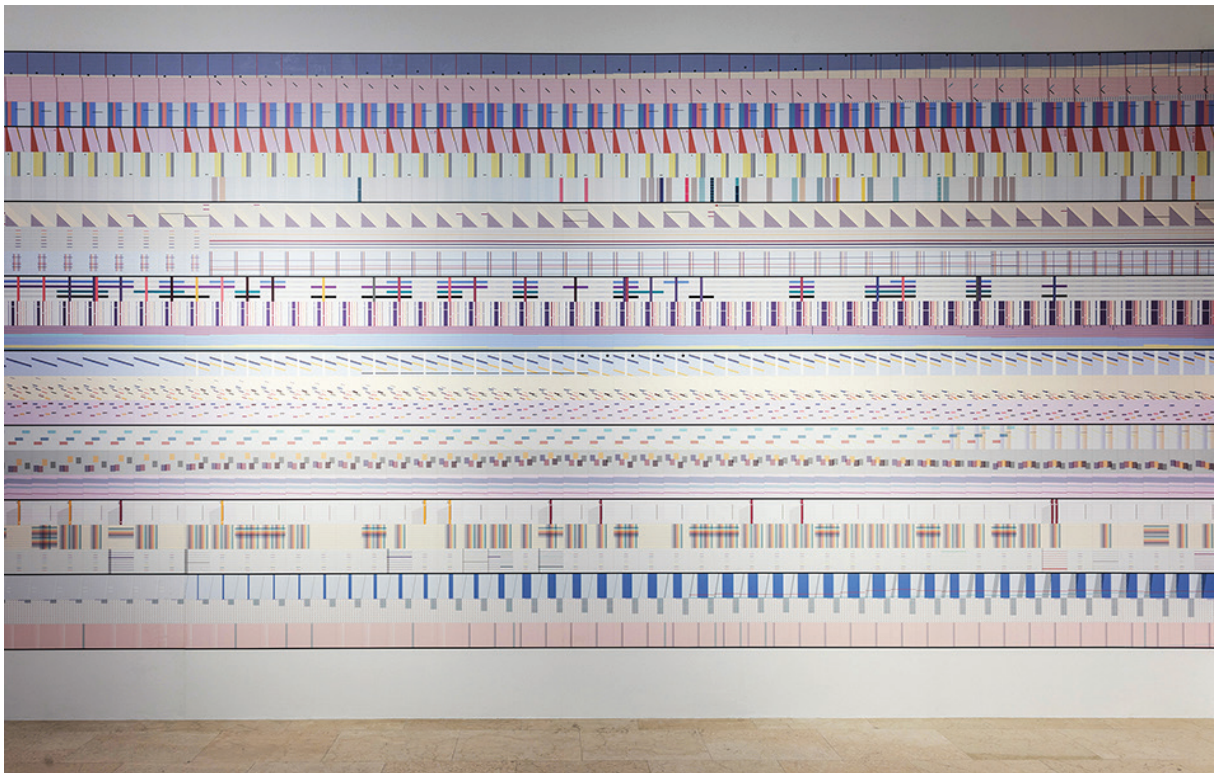

Figure 9. Accurat, *The Room of Change*, detail of the central wall
(© La Triennale di Milano; photo Gianluca Di Ioia).

The only exception are some color scales, which are always identical and used to represent the continents according to two different categorizations, one by geographic affinity and the other by cultural affinity (identical to the former but with the addition of another color for North Africa and the Middle East, an affinity that is clearly cultural and not geographic).

Thus, subject to the geographical references and their fixed color coordinates, a different regime of information visualization applies in each of the 24 datasets: from series to series, different plastic categories become relevant, and on all levels (chromatic, eidetic and topological). A curiously very nonfunctional visualization work, according to the parameters of classical data visualization, in which Tufte's data-ink ratio is completely disregarded. And to this great visual complexity it is difficult, based on this first description, to find an explanation according to the principle of transparency and perfect equivalence between data and its visualization.

Why all this complexity, then, in the visual transposition of information? To understand why, it is necessary to take a closer look at *The Room of Change* not only as an informational artifact, but as a complex visual artifact, as well as on its simultaneously informational and expressive nature, well summarized by the label of "data-driven wallpaper."

An initial observation can guide us in this direction. All the connections between the information and its graphic counterpart that we have just explored are in fact not available, as is good practice, on the same medium on which the visualization is inscribed. On closer inspection, on the large historiated wall of colored lines there is no title, nor are there any inscriptions, captions, or other indications.

The large and complex legend that holds everything one needs to decipher this data visualization–a sort of Rosetta Stone–on a cube in the center of the room (visible in Fig. 8), roughly at the point from which one can get a full and complete view of the three walls. After crossing the entrance line (the corridor between the two screens), and in any case already before arriving at the center, the visitor is invited to become an observer of a visual artifact, a complex and richly modulated abstract image. He/she is naturally invited to do so by the context (that of the Triennale Museum, and of the *Broken Nature* exhibition) as well as by what he is looking at.

After reaching the stele, however, in the center of the room, the observer acquires an awareness that he or she could not have before, upon entering: that of no longer being merely grappling with an aesthetic object, but with an informational object.

In classical information design, in fact, which Engelhardt (2003) explores, "reference objects" are always necessary for one to have an "informational object", that is, a well-regulated frame of reference is needed for specific visual elements to be understood as a data visualization, just as inscriptions, captions, numerical scales,

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

texts and other elements are needed that syncretically provide a unified whole with the visual object that is charged with conveying information.

It is therefore only upon reaching the stele, and reading the legend, that the visitor can realize that among the patterns modulating on the wall as in a tapestry there is not only something to *see*, but something to *know*. Thus, the level of pertinence of the viewer's gaze shifts, now mobilized to search for the information as it is being translated into colors and shapes: that is, those colors and shapes do not just compose a purely decorative panel, as it seems at first glance from the entrance corridor, but hold a message that needs to be deciphered.

In soliciting an active cognitive intervention on the part of the viewer, which is essential for the project to unfold, and in this its establishment through an enunciational device (the "activation" of the dataviz through the application of the legend), it seems to us that the visual complexity of the project finds its justification.

The visual density of the tapestry, in short, is not excessive or gratuitous. It exists to be discovered and interpreted by someone. But, at the same time, the real goal of such a project is not to offer a functional and systematic exploration of data, but to overwhelm the viewer/reader with an enormous amount of detailed information about climate change–and this information is conveyed in ways that are anything but automatic or immediate (see what has been said above about the proliferation of visual variables adopted).

In other words, *The Room of Change* exploits the non-immediate translatability of expression into content to solicit the viewer and transform him or her into the recipient of a specific discourse that first and foremost tells the millennial complexity of a very often underestimated issue.

The operation of deciphering the correspondences between visual elements and information on themes and macro-themes, then, which as we have seen is far from immediate, forces the viewer to become an attentive reader. The more he/she is involved, the more time he/she will spend trying hard to grasp as many narratives as possible about the many radical changes in our environment, from fossil fuel consumption to the effects of climate change. And the more he or she will want to know the more he or she, from a mere passing viewer in an exhibition, will become an informed reader, a counterpart to a very specific discourse: that of Accurat's data visualization and its ruled complexity.

Finally, it is the title of the installation (*The Room of Change*) that also alludes to a further, possible pathemic development of the viewer's gaze: that of a radical, real, and active change that comes through data knowledge and active participation. Indeed, the participation it takes to transform a giant tapestry into a data visualization is an excellent starting-point.

The last example we propose to consider is the data visualization published in the *New York Times*, February 21, 2021, on the million deaths from Covid-19 in the US.

Unlike the previous example, this is an extremely simple data visualization, based on essential visual means and structured in a rigorous manner. The graphic is simply composed by dots on a vertical timeline, and each of the nearly 500,000 individual dots represent a life lost in the United States to the Coronavirus. The points stretch chronologically down a long scroll, from the first reported U.S. death nearly a year ago to the current toll of often thousands of casualties per day.

From a purely informational point of view, from the perspective of the data-ink ratio, every bit of ink has a clear reason: each dot makes visible, as a single and perceivable entity, a life lost, with its position in the page layout representing the single day of the event.


The New York Times

U.S. VIRUS DEATHS NEARING 500,000 IN JUST ONE YEAR

MORE THAN IN 3 WARS

Empty Spaces in Cities, Towns, Restaurants, Homes and Hearts

The Toll: America Approaches Half a Million Covid Deaths

Garland Faces Resurgent Peril Of Extremism

Oklahoma City Attack Shaped His Views

STORMS EXPOSING A NATION PRIMED FOR CATASTROPHE

CLIMATE CHANGE WRATH

Unprepared for Threats Facing Power Grids, Water and Roads

Russia Fears But Can't Quit Open Internet

Lack of Birth Control Deepens Women's Burden in Venezuela

Israel Funds Vaccines for Syria

Osaka Wins Australian Open

Sit. Rail Over. Wait Your Turn.

Boredom Is Making Us Buy

Katie Engelhart

Figure 10. Front page of Sunday's *New York Times*, February 21, 2021, graphics by Lauren Leatherby.

Quantitatively, there is no diagrammatic reduction of data (e.g. pictograms, in Neurath's Isotype style, or simple geometric shapes standing in for a quantity of objects or events): all individual deaths are equally represented. There is, instead, a reduction of information about each event, which effectively flattens the death of the individual to its irreducible human singularity, to the temporal coordinates of the event alone.

This is a very different choice from the one made for the May 24, 2020 front page (fig. 11), which reported, of the first 100,000 deaths on U.S. soil, a sample of 1,000 deaths displayed as a lists of names and obituaries, choosing to shape the same event (deaths in the U.S. for Covid-19) with decidedly more personal connotations. Simone Landon, assistant editor of the Graphics desk, said the aim was "to represent the number in a way that conveyed both the vastness and the variety of lives lost," in fact avoiding reducing people to pure dots or stick figures but choosing to tell something "about who these people were, the lives that they lived, what it means for us as a country" (Grippe 2020).

Punctum. International Journal of Semiotics | **08:01:2022**
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

"All the News That's Fit to Print"

The New York Times

Late Edition

VOL. CLXIX ... No. 58,703 NEW YORK, SUNDAY, MAY 24, 2020 $6.00

U.S. DEATHS NEAR 100,000, AN INCALCULABLE LOSS

They Were Not Simply Names on a List. They Were Us.

Numbers alone cannot possibly measure the impact of the coronavirus on America, whether it is the number of patients treated, jobs interrupted or lives cut short. As the country nears a grim milestone of 100,000 deaths attributed to the virus, we have asked a few wonderful souvenirs to list hundreds of the victims. The 1,000 people here reflect just 1 percent of the toll. None were mere numbers.

Figure 11. Front page of Sunday's *New York Times*, May 23, 2020, graphics by Simone Landon.

Thus, from the May 24, 2020 visualization to the February 21, 2021 visualization, both the extent of the reference dataset (from 100,000 to 500,000 data points) and the way in which data is designed have changed. What was previously described, for example, as "Theresa Elloie, 63, New Orleans, renowned for her business making detailed pins and corsages ..." (Grippe 2020) is now nothing more than an anonymous black dot, with no more name, age, or biographical profile, in a graphic "with nearly a half-million dots running down the length of the page and across three of its six columns" (Coleman 2021).

This is a radical change that *New York Times* readers, in comments on the Facebook post announcing the new front page on February 21, 2021, did not fail to remark on, with comments such as "So my father in-law is basically a dot to some. Horrible way to describe someone's life" or "Herbert Shigekane is one of the dots... thank you for honoring the dead." As also in the cases examined above, the application of categories (and filters) for data collection turns out to be a true display device, capable of also informing the subsequent visual return of the data.

As for the visual display strategies of the specific data, although that of February, 21 is a purely diagrammatic visualization, with very few visual variables at play, it would be incomplete to describe it only from an informational point of view. We should not forget the expressive power that this visualization strategy of deaths for Covid-19 brings, highlighting the dizzying rapidity of the epidemic even in a very short time, as time goes on, and the consequent increase in deaths in some periods, particularly in the second part of the visualization (with a peak of more than 50000 deaths in 17 days, from January, 2 to January, 19).

Indeed, more than the individual dots, which are still clearly identifiable even if no longer narrated as in the previous visualization, the observer's eye is struck by the horizontal bands in which, as in a typographic screen, the blackness thickens without any form –- any data–being clearly distinguishable.

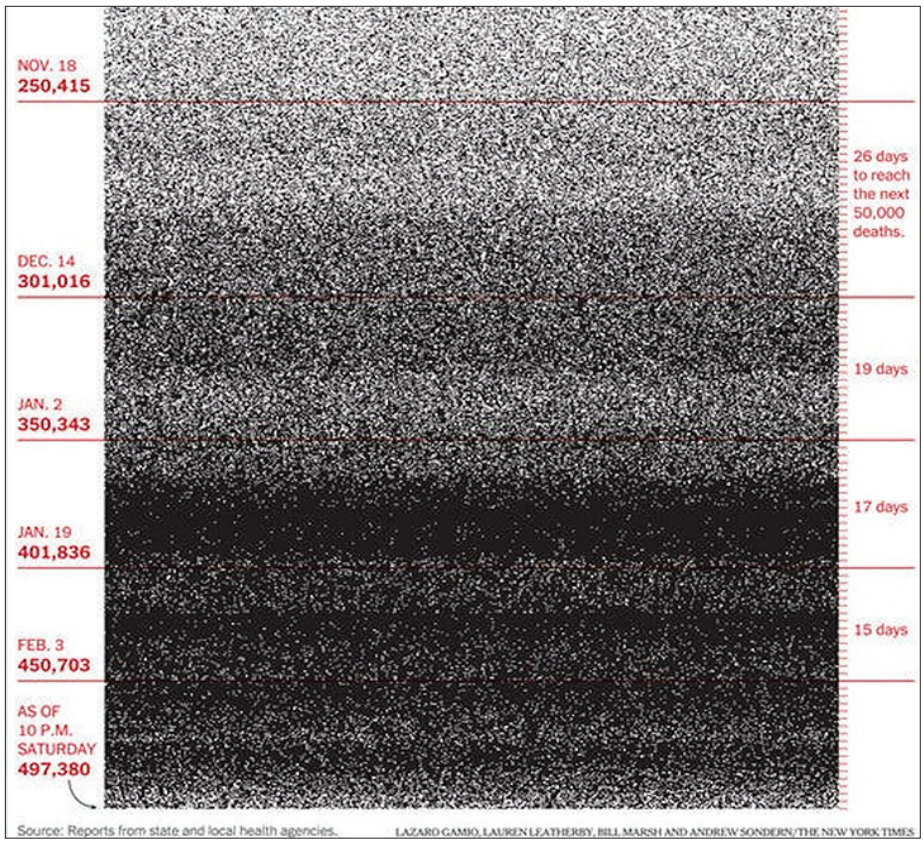

Figure 12. Front page of Sunday's *New York Times*, February 21, 2021, graphics by Lauren Leatherby, detail.

The purely functional strategy adopted in the visual representation of data fades into the tragic impossibility of quantifying such a large number in such a short period, as if to remark on the fallibility of exact, live recounting of massive amounts of data, and at the same time the need we still have to make the invisible visible.

In this case, the visualization of such a large database, precisely because of the arduous task it sets itself, proves capable of making one feel rather than see, transforming the information receiver into an observer and a pathemic subject as well. The reader, indeed, guided in the description of the phenomenon by the temporal legend running to the left and by the comments running to the right of the visualization marking the days intervals, becomes progressively immersed in an increasingly dense flow of data. Moreover, he/she is more and more gripped by the tense crescendo that opposes, semi-symbolically, the plastic discontinuity between the dots in the white field (from the top up to 4/5 of the image), which represent a pandemic situation that is still tellable, still numerable, still discrete, to the increasing density of the dots in the black field (bottom), mere plastic continuity in which no distinctions can be made or data identified, in which it is almost impossible to distinguish anything.

Punctum. International Journal of Semiotics | **08**:01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

Paradoxically, then, what seemingly appears to be a purely functionalist visualization geared toward the pure display of data, with no other visual connotations (we are a long way from the language of infographics), turns out instead to be a tool for making the inconceivable tangible.

A veritable data-wall (a "wall of grief," it has been called), which brings attention once again to the impossibility of a purely immediate or transparent visualization in its giving access to data, and which shows instead how any strategy of putting data into discourse, even the most functionalist in means and intent, is but an attempt to make something otherwise invisible visible.

5. Across immediacy and transparency. Conclusions

In this paper we have tried to consider visualization no longer as a mere display of data but as the result of a multi-level translation process, from structuring data into datasets to visually transposing a specific viewpoint on the data. As we have tried to show, such a semiotic exploration, attentive to the ways in which data are put into discourse, could work as a deeper and more structural investigation of the social, cultural and at the same time discursive construction implicit in the big data phenomenon.

As for media visualization examples, in *Visualizing Vertov* we have remarked the crucial role of metadata and different combinations of visual characteristics in informing both media visualizations and patterns, as well as in *On Broadway* we have noticed how the quite "immersive" effect of a bottom-up perspective in the so-called Streetview Top layer depends on the capillary organization of data collection, a real "device of visualization."

Regarding the data visualization projects analyzed, in *The Room of Change* we have observed how an abstract visualization can be read either as a purely aesthetic object or as an informational object, depending on the level of relevance of the viewer's gaze, and how the shift from an aesthetic reading to an informational reading can generate a pathemic engagement in the viewer.

Finally, on the front page of the *New York Times* of Sunday 21 February 2021, we noted how even a purely functional and seemingly immediate visualization of a large amount of data can also generate, at the same time, an entirely different perspective on the data displayed.

Manovich is very clear in pointing out that direct visualization is also a heuristic method, allowing the exploration of large collections of data that is difficult to manage in other ways. The role of diagrams, as well of visualizations in shaping knowledge and not just in representing it, is well known (see for example Netz 2003, Stjernfelt 2007, Thurlemann 2008).

An exploratory approach to data can indeed be found not only in direct visualization but also in many examples of data visualization, offering not just a "portrait" of a phenomenon but a map of virtualities to be activated and explored, as in Accurat's *Room of Change* as well as in other more standard projects managing data complexity. In any case, as we have seen, we can focus on such objects not just as pure systems of virtual features ready for future operations but as texts, with their specific visual communication strategies.

In particular, we have tried to show how a visualization that returns a large amount of data in a relatively visually simple way, such as the front page of Sunday's *New York Times*, February 21, 2021, is not only an exploratory and purely informational tool but can also, and at the same time, produce a strong pathemic effect in the reader/viewer.

To return again to Françoise Bastide and her description of scientific image (and data image, we would add) as a chain of devices of visualization, it seems that the two Covid-19 *New York Times* visualizations show how the choice of a different categorization of the same data (narrative-qualitative, in May 24, 2020 vs. quantitative, in February 21, 2021) and a different visual translation of the data (a sample of names and biography of the deaths in the 2020 visualization vs. the totality of the deaths represented by dots in the 2021 visualization) generate an entirely different meaning-effect.

In the first case, we are faced with an almost point-to-point mapping of a sample of stories and biographies of the protagonists of the Covid-19 tragedy, which we are free to explore and read as we wish. In the second case, the effect is that of a top-down view that, in order to represent reality as accurately and abstractly as possible, exposes the viewer to the risk of seeing everything without being able to rationally understand what he/she sees.

Ultimately, it seems to emerge, in a clear way, how *making known* (in data visualization, through the organization and interpretation of data from a given perspective; in media visualization, through the various dimensions of image analysis) is an ineradicable component of *making visible*, and is proper to any visualization, as well as to the rhetorical, argumentative, and discursive strategies that make it possible.

Punctum. International Journal of Semiotics | **08**:01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

References

Bastide, Françoise 1981. La démonstration, analyse de la structure actantielle du faire-croire, *Actes sémiotiques. Documents du Groupe de Recherches sémio-linguistique*, III (28): 7-8. Translation: On Demonstration. An Analysis of the Actantial Structure of Causing-to-Believe. In: Perron, Paul and Collins, Frank (eds.) 1989, *Paris School Semiotics. II. Practice*, Amsterdam/Philadelphia: John Benjamins, 109-143.

Bastide, Françoise 1985a. Essai d'épistemologie à partir d'un texte technique sans prétention: une invention peu connue des frères Lumière, *Fundamenta scientiae*, VI, (2): 127–150.

Bastide, Françoise 1985b. Iconographie des textes scientifiques: principes d'analyse, *Culture Technique* 14: 133–151. Translation: The iconography of scientific texts. In: Lynch, Michael and Steve Woolgar (eds.) 1990, *Representation in scientific practice*, Cambridge, Mass.: MIT Press, 187-229.

Bastide, Françoise 2001. Bruno Latour and Paolo Fabbri (eds.) *Una notte con Saturno. Scritti semiotici sul discorso scientifico*, Milano: Meltemi.

Bertin, Jacques 1967. *Sémiologie graphique. Les diagrammes, les réseaux, les cartes*, Paris-La Haye: Gauthier-Villars/Mouton & Cie.

Boyd, Danah, and Crawford, Kate 2012. Critical questions for big data. Provocations for a cultural, technological, and scholarly phenomenon, *Information, Communication & Society* 15 (5): 662-679.

Card, Stuart K., J.D. Mackinlay, and B. Shneiderman (eds.) 1999. *Readings in Information visualization. Using vision to think*. San Francisco: Morgan Stanley Kaufmann Publishers.

Coleman, Nancy 2021. On the Front Page, a Wall of Grief, *The New York Times*, Feb 21. https://www.nytimes.com/2021/02/21/insider/covid-500k-front-page.html>.

Corrain, Lucia and Tarcisio Lancioni 2012. Geometrie del senso. Ripensare la semiotica dell'immagine a partire da *La macchina della pittura* di Omar Calabrese. E|C, 1 (12). www.ec-aiss.it/includes/tng/pub/tNG_download4.php?record ID=657&KT_download1=a6512c6693a3506026073b273a6ec755.

Crawford, Kate 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven and London: Yale University Press.

Dondero, Maria Giulia, and Fontanille, Jacques 2012. *Des images à problèmes. Le sens du visuel à l'épreuve de l'image scientifique*. Limoges: PULIM. Translation: 2014. *The Semiotic Challenge of Scientific Images*. A Test Case for Visual Meaning. Ottawa: Legas Publishing.

Dondero, Maria Giulia 2017. The semiotics of design in media visualization. *Information Design Journal* 23(2): 208-218.

Dondero, Maria Giulia 2018. La remédiation d'archives visuelles en vue de nouvelles iconographies: le cas de la Media Visualization de Lev Manovich. *Interin* 1(23): 85–107. https://seer.utp.br/index.php/i/issue/view/41.

Dondero, Maria Giulia 2020. *The Language of Images: The Forms and the Forces*. Dordrecht: Springer Netherlands.

Drucker, Johanna. 2020. *Visualization and interpretation. Humanistic approaches to display.* Cambridge, Mass.: MIT Press.

Fabbri, Paolo 1998a. *La svolta semiotica*. Roma-Bari: Laterza.

Fabbri, Paolo 1998b. L'oscuro principe spinozista. Deleuze Hjelmslev Bacon, *Discipline filosofiche* 1: 209-220.

Fabbri, Paolo 2001. Introduction to Bastide, Françoise 2001. Bruno Latour and Paolo Fabbri (eds.) *Una notte con Saturno. Scritti semiotici sul discorso scientifico*, Milano: Meltemi, 9-23.

Fabbri, Paolo 2014. Diagrammi in filosofia. G. Deleuze e la semiotica "pura". In: V. Mancia (ed.) *Carte Semiotiche. Immagini che fanno segno*, 27-35.

Fabbri, Paolo 2021. *Rigore e immaginazione. Percorsi semiotici sulle scienze*, P. Donghi (ed.). Milano: Mimesis.

Floch, Jean-Marie 1985. *Petites mythologies de l'œil et de l'esprit. Pour une sémiotique plastique*. Paris & Amsterdam: Hadès/Benjamins.

Giest, Sarah, and Samuels, Annemarie 2020). "For good measure": data gaps in a big data world. *Policy Sci* 53: 559-569. <https://doi.org/10.1007/s11077-020-09384-1>.

Greimas, Algirdas Julien 1984. Sémiotique figurative et sémiotique plastique. *Actes Sémiotiques*, 60: 5-24.

Grippe, John 2020. The Project Behind a Front Page Full of Names. The New York Times, May 23. <https://www.nytimes.com/2020/05/23/reader-center/coronavirus-new-york-times-front-page.html>.

Hand, David J. 2020. *Dark data. Why what you don't know matters*. Princeton: Princeton University Press.

Lancioni, Tarcisio 2013. Immagini in prospettiva. Forme e figure dell'enunciazione visiva. In: Massimo Leone and Isabella Pezzini (eds.) *Semiotica delle soggettività. Per Omar*. Roma: Aracne, 281-288.

Latour, Bruno. 1987. *Science in Action. How to Follow Scientists and Engineers through Society*. Cambridge Mass: Harvard University Press.

Latour, Bruno. 2009. La sémiotique des textes scientifiques depuis le travail de Françoise Bastide. *Visible* 5: 251-262.

Latour, Bruno and Bastide, Françoise 1986. Writing Science. Fact and Fiction. In: Michel Callon, John Law, Arie Rip (eds.), *Mapping the Dynamics of Science and Technology*. London: Macmillan, 51-66.

Latour, Bruno and Fabbri, Paolo 1977. "La rhétorique de la science. Pouvoir et devoir dans un article de science exacte". *Actes de la Recherche en sciences sociales.* 13(1): 81-95. Paris: Minuit. Translation: 2000. The Rhetoric of Science: Authority and Duty in an Article from the Exact Sciences, *Technostyle* 16 (1): 115–134.

Punctum. International Journal of Semiotics | **08:**01:2022
ISSN 2459-2943 | DOI: 10.18680/hss.2022.0006 | **punctum.gr**

Latour, Bruno and Woolgar, Steve 1979. *Laboratory Life. The Construction of Scientific Facts*. Los Angeles: Sage Publications. Revised edn. *Laboratory Life: the Construction of Scientific Facts*, Princeton, Princeton University Press, 1986.

Lerman, Jonas 2013. Big data and its exclusions. *Stanford Law Review Online* 66: 55-63.

Manchia, Valentina 2014. Cinema Under the Microscope. Cultural Analytics and Data Visualization for Media Studies: the Visualizing Vertov Project. *Progetto Grafico*, 25: 68-79.

Manchia, Valentina 2020a. *Il discorso dei dati. Note semiotiche sulla visualizzazione delle informazioni*. Milano: FrancoAngeli.

Manchia, Valentina 2020b. Dati, interattività, immersione. Alcune note tra data visualization e media visualization. *E/C*. 30: 374-383.

Manchia, Valentina 2021. Dati mancanti, dati mancati. Cancel culture, data bias e data gap nell'era dei big data. *Filosofi(e)Semiotiche* 8(1): 62–73.

Manovich, Lev 2010. What is visualization?, <http://manovich.net/content/04-projects/064-what-is-visualization/61_article_2010.pdf>. Now republished in Manovich 2020.

Manovich, Lev 2013. Visualizing Vertov, http://softwarestudies.com/cultural_analytics/Manovich. Visualizing_Vertov.2013.pdf.

Manovich, Lev 2020. *Cultural analytics*. Cambridge, Mass.: MIT Press.

Marin, Louis 1983. La ville dans sa carte et son portrait : proposition de recherche. In: *De la représentation*. Paris: Seuil-Gallimard, 204-218.

Mengoni, Angela 2021. Visual Semiotics. In: Krešimir Purgar (ed.), *The Palgrave Hand book of Image Studies*, pp. 41-654. https://doi.org/10.1007/978-3-030-71830-5_39

Netz, Reviel 2003. *The Shaping of Deduction in Greek Mathematics. A Study in Cognitive History*. Cambridge: Cambridge University Press.

Olteanu, Alexandra, Castillo, Carlos, Diaz, Fernando, and Kiciman, Emre 2019. Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Front. Big Data* 2:13.

Peirce, Charles S. 1931-1935. *Collected Papers*. Cambridge, Mass.: Harvard University Press.

Peverini, Paolo 2019. *Alla ricerca del senso. Bruno Latour in dialogo con la semiotica*. Roma: Edizioni Nuova Cultura.

Stjernfelt, Frederik 2007. *Diagrammatology. An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics*. Dordrecht: Springer Netherlands.

Thürlemann, Félix 2008. Diagramme ou regard diagrammatique?, *Visible* 4: 137-150.

Tufte, Edward R. 1983. *The visual display of quantitative information*. Cheshire, Conn.: Graphics Press.

Tufte, Edward R. 1990. *Envisioning Information*. Cheshire, Conn.: Graphics Press.

Tufte, Edward R. 1997. *Visual Explanations. Images and Quantities, Evidence and Narrative*. Cheshire, Conn.: Graphics Press.

Tufte, Edward R. 2001. *The Visual Display of Quantitative Information*. Cheshire, Conn.: Graphics Press.

Warde, Beatrice 1930. The crystal goblet, or why printing should be invisible. In: *The crystal goblet*. London, Sylvan Press, 1955.

Zingale, Salvatore 2012. *Interpretazione e progetto*. Milano: Franco Angeli.

Zingale, Salvatore 2016. Design as translation activity: a semiotic overview. In: Lloyd, P. and Bohemia, E. (eds.), *Future Focused Thinking-DRS International Conference 2016*, 27-30 June, Brighton, United Kingdom.

AUTHOR

Valentina Manchia Adjunct Professor at the Politecnico di Milano, Alma Mater Studiorum–University of Bologna, and ISIA Urbino and a member of CROSS Research Centre.