

The “modulation mania”

Impersonal enunciation and visual rhetoric of the data image

punctum.gr

BY: Federico Biggio

ABSTRACT

The growth of contemporary positivistic trust in computational tools inspired semiotics to contribute to the debate on the strategies of communication of truth. In order to study these strategies in data images, the article features a framework modelled on the theory of impersonal enunciation. The computational utterer is believed to assure a super-personal enunciation, understood as impartial and true. Furthermore, it will be observed how subjectivities are in any case modulated in the data images: this operation constitutes as much a visual rhetoric to achieve the effect of truthfulness as it is a discursive practice featuring the computational enunciation of the data image.

ARTICLE INFO:

Volume: 08

Issue: 01

Summer 2022

ISSN: 2459-2943

DOI: 10.18680/hss.2022.0003

Pages: 31-45

Lic.: CC BY-NC-ND 4.0

KEYWORDS:

Computational enunciation

Truthfulness

Data Scraping

Data Images

Data pollution

1. The framing of truthfulness

Today, data visualisation, data-scraping, computational and automated data analysis software can be understood as techno-linguistic artefacts that users employ in order to gain an interpretative advantage. The hypothesis of this article is that semiotics can better address the epistemic challenge posed by data analysis with which scientific communities are grappling today.

One of the main themes of contemporary thinking on this topic concerns the effects of meaning produced by data visualisation images, or data images, and since semiotics has established itself as the science of texts and meaning, the semiotics of big data often focuses on these kinds of visual objects (Reyes 2017; Dondero 2020).

To this end, it is important to note that, if we consider the data image as the result of a computational enunciation, it is possible to argue that a certain effect of truthfulness (Greimas 1984)—which is traceable to that of referential illusion or objectivity (Greimas & Courtés 1979: 116)—is closely linked to the supra-personal and encyclopaedic nature of the computational enunciator. If we assume the existence of such effects of meaning, it is possible to assert that the data image is a textual form, the result of a predicative practice aimed at obtaining an interpretive gain for the human utterer who once a data image has been realised and enunciated, can be configured as a communicative gain.

In this perspective, every data image is the result of a computational enunciative practice (Fontanille 1994). A data analysis session and the enunciation of a data image are guided by the collaboration between a user and a computational entity. An example of this enunciative process is found in the collection of data disseminated over a network. To extract a cluster from the data, the collector manipulates the discontinuity of the data streams. Through the collection process, a set of data is extracted from databases and organised into clusters and features to provide a meaningful reading of a process. Through a Google Trends search, users can understand what are the most important “keywords” mentioned by other users in their tweets, posts, or search engine queries. During a data analysis, data generated during interactions with previous users may have been identified and grouped into semantic categories based on which the elements of the data image were organised and modelled.

In trend analysis software, users can retrieve a series of “strings” (referring to the interactional-spatial utterances of other users), process them, “clean them up”, and visualise them in a data image. Such computational utterances are “meta-communications” in that users produce an interactional-spatial utterance to communicate a representation of a set of discourses about human-computer interaction of *other* users and semantically organise them into behavioural categories. For example, these operations may consist of simultaneously viewing images produced by users and shared on Instagram. Trend analysis software, on the other hand, allows users to manipulate a data set more quickly and (seemingly) more fairly than the user could. At the heart of these operations is the definition of contrasts and divergences, boundaries and attributes, semantic categories; it is the automatic assignment of specific features, labels, and abstract metadata using programming metalanguage to achieve interpretive gain.

Such an enunciative practise is able to describe the values of objectivity and truthfulness in the data image, thus to achieve a gain for the users. This idea is prevalent in computer and media science.

For some, at the ontological level, data are benign. Data are simply data, essential elements that are abstracted from the world in neutral and objective ways subject to technical constraints. [...] They are pre-analytical and pre-factual. From this perspective, a sensor has no politics or agenda. It simply measures light or heat or humidity, and so on—producing readings that reflect the truth about the world unless tainted by

a technical glitch. In other words, the sensor produces an objective, realist view of the world revealing things as they really are, wherein the reality of thing being measured is independent of the measuring process (Kitchin 2014: 48).

Assuming the commonly shared belief in computing, the first aspect to consider is the artificial nature of the computer enunciator. In this context, drawing on the framework of impersonal enunciation elaborated by Claudio Paolucci (2020), one can formulate the idea that the interpretive gain of data analysis and visualisation operations lies in the—alleged—impersonality of the computational utterer. Impersonality is defined by Paolucci as a property of the non-human actor, as is the computational utterer.

In this context, a cultural analysis should consider the innate fascination with predictability and computational predictions of the future that has appealed to humans since the dawn of time, sometimes in search of universal equilibrium (Cramer 2005; Pasquinelli 2019). In this perspective, “super-personality” can be assumed to be a feature of computer languages, which are metalanguages (Sieckenius de Souza 2005) that can guarantee their meta-pragmatic functioning, just as the meta-communications of the sciences do, through which the speakers of a cultural community implicitly or explicitly establish the rules of referentiality and the values of objectivity.

However, the non-human actor is not the one who produces the empirical utterance. This is done through the human and non-human pair of utterances: “if we consider the same opposition human VS non-human, we can very well see that it is not a “non-person” (non-human) at all, but rather a “person + non-person” (human + non-human)” (Paolucci 2020: 67).

On the one hand, it can be said that this gain is interpretative in nature, as data processing algorithms perform computational analysis more accurately and correctly than a human, and for this reason provide a gain to users seeking a quantitative result in their search. A data image that graphically represents the data extracted from the acquisition processes brings some gain in that it allows one to move from the indeterminate impure to the determinate pure, thus leading to the acquisition of objective knowledge (and thus an effect of truthfulness) that comes from the impersonality of the computational calculation. In this view, it is not the algorithms that want something from us (as Finn noted in 2017), but the users that want something from the algorithms. This interpretive gain with data analytics software is closely linked to the establishment of a “distant gaze” that users adopt in their interactive-spatial expressions. It is precisely the process of moving away from the subjective, embodied, and situated scopic regime that allows us to formally speak of super-personality.¹

¹ The distant reading paradigm, developed by Franco Moretti (2005) in the field of literary criticism denotes a method of computational analysis of written texts belonging to a single literary current or referable to a single author (but the same can be done with visual texts as Manovich’s cultural analytics does). This method is able to bring out trans-textual isotopies and recurrent stylistic elements into a corpus of texts: “distant reading: where distance, is a condition of knowledge: it allows to focus on units that are much smaller or much larger than the text: devices, themes, tropes or genres and systems” (Moretti 2005: 48-49).

On the other hand, as we shall see, it is not only the computational calculation of the artificial exterior that produces the gain in truthfulness and objectivity. Although the co-enunciative practice of the data image suppresses the personality and subjectivity of the human enunciator, one can observe in several cases how a new form of subjectivity, plural and collective, emerges. It is this latter form of subjectivity that characterizes the encyclopaedia and the quantitative determination of truth in data images.

Based on these ideas, this gain also lies in the ability of data images to self-reflect users and their mutual relations, to give them tools (images) to understand themselves.

From this perspective, the main property of data images lies in their explanatory potential and their ability to make visible relationships and dynamics that are invisible to the naked eye. An example of this type of practice is data analysis in socio-geographic research, which is conducted to self-reflexively understand the cultural dynamics that characterize a sociological context. In a 2015 study, Marta Severo compared several national networks (France, Italy, and Switzerland) involved in the implementation of the 2003 Convention UNESCO to highlight national trends and specificities, using online “traces” to study offline phenomena such as routines, oral traditions, and cultural practices. The analysis resulted in image data and, in effect, heat maps depicting clusters of topics in relation to the social position of the pronouncing actor.

Much like the Internet, the Web is not a single network but a network of networks, a graph in which densely connected zones are separated by relatively empty spaces. For the most part, these zones correspond to thematic communities: groups of people clustered around interests and points of view. The hyperlinks between sites dealing with human-computer interaction can therefore reveal much about the networks of actors involved in the conservation of human-computer interaction. In other words: If we know which websites are linked online, we can infer which actors are linked offline (Severo & Venturini 2015: 4).

Here, we can clearly observe the meaning that comes from following hyperlinks along users’ paths and analysing digital traces. Moreover, this concept of “reflexivity” implies certain practices that are closely related to the creation of an effect of transparency.

The analysis confirmed that national and institutional nodes (especially the overlap of the two categories, i.e., national institutions) play the role of authorities in the networks, as they receive the most links and are consequently very visible sites in the corpus. [...] While national institutions occupy the centre of the networks, local and individual nodes tend to be located at the edges of the graphs, and local sites are often densely clustered around specific cultural practices. Moreover, as expected, the network structure of each country reflects the main characteristics of its cultural system (Severo & Venturini 2015: 15).

The data meditation rituals of Salvatore Iaconesi and Oriana Persico (2020), concerning autobiographical and collective writing paths that were graphically organised

into data images and multisensory images with strong evocative value using data visualisation software, and also used for meditative purposes, were also in this direction.

In such a view, to be reflected in the data image is the attitude of a social actor, which is a collective and a plural one. The lesson we can draw from the impersonal utterance, then, is that it is characteristic not only of a particular type of discourse (the free indirect speech) or of a social value (the a-personality), but also of an “inner” discourse, of an “I” saying “I” (Paolucci 2020) and opening up to self-awareness.

2. Impersonality, encyclopaedia and database

According to Paolucci, “impersonality” describes the determination of culturally established (i.e., encyclopaedic) norms, schemes, and customs on subjective (i.e., personal) enunciates. In this context, Paolucci cites Guillaume and “the primacy of he/she” in describing impersonality as the quality of the delocutive form as an “evenemential subject that opens up subject positions in which *one* speaks” (Paolucci 2020: 73). In such a view, the theory of impersonality not only allows us to describe the general functioning of enunciation, but also to affirm a super-personality at the level of language, which ultimately allows us to describe cultural stereotypes, clichés, and genres.

In such a perspective, enunciation is impersonal in so far as to produce an utterance is to *keep in the present* a set of norms, customs, and differential relations that constitute a schema (Eco’s encyclopaedia), i.e., to adhere to a certain *implicit operative context* that is meta communicatively presupposed. In such a perspective, the epistemic tradition of Benveniste is rejected, according to which the subject of the enunciation selects the linguistic elements of the *langue* in order to realize them in the *parole*.

Although Eco’s conception of the encyclopaedia was mainly a theoretical and abstract one that could refer to a specific typology of knowledge located before the spread of the Web—and although the encyclopaedia is theoretically the totality of knowledge and information, whereas the Web is just one of the archives of knowledge and information—it could have great heuristic value to find in it a theoretical place or level to explain “digital crumbs” (Pentland 2014) and “digital traces” (Severo & Romele 2015; Ferraris 2021). Indeed, we now know that the distribution of online media content must be considered against the background of a complex, deep information system that determines the order in which content appears, a system that has been called “web platform.” This system is typical of social media or search engine feeds; it is based on algorithmic calculations of data about the user’s online activities. It constantly creates digital simulacra of the world and directly influences the perception of the world. One could say that it is the production of an ekphrastic mathematical representation of the existence of individuals who extend themselves and produce data through their actions in the digital space.

Social physics functions by analysing patterns of human experience and idea exchange within the digital bread crumbs we all leave behind us as we move through the world—call records, credit card transactions, and GPS location fixes, among others. These data tell the story of everyday life by recording what each of us has chosen to do. And this is very different from what is put on Facebook; postings on Facebook are what people choose to tell each other, edited according to the standards of the day. Who we actually are is more accurately determined by where we spend our time and which things we buy, not just by what we say we do (Pentland 2014: 22).

In this perspective, recovering Umberto Eco's (1979; 2007) concept of the encyclopaedia will be useful to define a paradigmatic semantic horizon that brings together all the knowledge of a community and the information produced during a practice of human-computer interaction with digital media. Consequently, the *langue* of computer utterances is precisely the set of digital archives (databases) linked to special software capable of scraping such encyclopaedic spaces. Also, the utterances provided by the methodological-applicative proposals of data science can be called "enunciation of the encyclopaedia", i.e., utterances that access the archive of a large number of differentiated utterances and make it possible to draw continuity inferences and derive semantic isotopies.

If we accept the definition of the encyclopaedia as "the totality of all utterances", the set of "what has already been said", we could consider the database as the totality of all digital utterances that have already been announced and translated into data. According to Lev Manovich, one of the most important theorists of software culture, the database is—to quote Panofsky—a "symbolic form", a structured collection of data. In his most famous book, *The Language of New Media* (2001), he argued that the computerization of culture is accompanied by the advent of "storage mania", data collection, and indexing. Such a process is not only a linguistic operation aimed at producing a data image, but also a practice charged with social, economic, and political values.

In this respect, we are at the beginning of a new education in which the digital (the computational, the logical-mathematical translation of the world into discrete units) constitutes the process of universal construction of knowledge (like the university pedagogy of education) (Berry 2011).

3. Impersonality, empiricism and narrativity

To deal with the gains from the social use of data analysis tools, the absence of the possibility of attributing personhood to the non-human actors who rule over the enunciation of feeds and data images can be described as both *impersonality* and *a-personality*.

According to Eugeni, "the first way is that of impersonality: according to this view, the drying up of the category of person leads to a homogenization of humans, animals, and machines, and thus to the so-called post-human" (Eugeni 2021: 212). This position

is retrieved by Paolucci, who, however, is not concerned with “post-human” environments as described by Eugeni (2021), but rather with the gain of impersonal tools.

For this reason, the enunciation in the realm of Big Data not only retains the “memory” of its future re-enunciation, but it is entrusted with a set of norms and uses that attend to the data, linking its semiotic format to a set of other social institutions whose particularities and norms (political, economic, ethical, etc.) influence the data journey itself (Paolucci 2020: 164-165).

On the other hand, Eugeni postulated the concept of *a-personality*: according to him, de-personalization (or delocution) is only a penultimate truth and considers it as a *pars destruens* that precedes a re-personalization: “an investigation of the forces in the field, finally free from imaginary and ideological prejudices, allowing us to understand, beyond the attempts to personalise artificial intelligences, what they actually do in the shared environment and who is actually responsible for their design and operation” (Eugeni 2021: 212).

Thus, it could be argued that the truth effect is determined by the utterance of a “super-personal” and computational utterer. It is a personal type of utterance, since this utterer is an artificial, non-personal entity. At the same time, it is an impersonal type of utterance, since the utterance contains several encyclopaedic dimensions of utterance, and it is this plurality and multiplicity that suppresses the subjectivity of the individual in rhetoric and makes the impersonal utterance appear as not characterised by the encyclopaedic features of its utterers.

The connection between the impersonal proclamation and the effect of truthfulness, however, might seem to philosophers of science to be a contradiction, and in part it is. As impersonal proclamation refers to the encyclopaedia, all data enunciations are “steeped” in culture and bias. As the majority of authors in the field of data studies note (Chun et al., 2019; Cairo 2019), “information visualizations are interpretive acts masquerading as presentations” (Drucker 2014: 10). On the contrary, according to Lotman (1985), cultures and their (technical) artifacts follow homeostatic mechanisms that ultimately assure people of the coherence and validity of circulating semantic content and beliefs.

Therefore, it can be argued that the impersonal enunciation of the computational actor produces a truth effect that is acknowledged as trustworthy.

In such a perspective, then, we need to consider the empirical utterer, and thus the social actors, who strategically deploy such an impersonal artifact to support a particular kind of enunciate. The functioning of the impersonal utterance would have to be considered in light of the rhetorical and narrative strategies employed. One solution might be to draw on the narrative paradigm of semiotics (Greimas 1984) to explain the ways in which designers manipulate data and information by articulating in linear and selective ways something that is not originally defined. David Bihanic (2015) argued that data design is the most important practice that needs to be explored and developed to address today’s “information overload”, and in the terminology of semiotics, this means exploring the discursive dimension of Big Data.

Narratives are not only one of the preferred objects of analysis in text semiotics; they are also and above all a set of models used by people to give meaning to things and experiences. Therefore, the narrators' perspective could be useful for the analysis of the implicit practices that lead an empirical subject to express a certain data image.

However, it was precisely Lev Manovich (2001: 199) to argue that:

As a cultural form, database represents the world as a list of items and it refuses to order this list. In contrast, a narrative creates a cause-and-effect trajectory of seemingly unordered items (events). Therefore, database and narrative are natural enemies. Competing for the same territory of human culture, each claims an exclusive right to make meaning out of the world.

After twenty-one years, however, his ideas about game practices—which he described precisely as narrative and interactive practices like those of data analysis—could maybe be revisited and reformed to explain narratively both the syntax of designers' practices and the interpretations of recipients.²

While such an application of the narrative paradigm may appear to be a method of deductive interpretation, it is nonetheless useful in focusing on the internal narrative opened up by data imagery. The provocative article by Willers (2015) entitled "Show, do not tell" aimed precisely at noting the "narrative approach [of data imagery] that presents the audience with a linear path to follow, dictated entirely by the author" (ibid: 4).

As far as impersonal enunciation is concerned, introducing the topic of narrative and discursive strategies of data images leads us to consider the practice of *modulating subjectivities*, which seems to be central to the process of constituting the objectivity effect.

In this sense, it is not only the artificial nature of the computational enunciator that is of central importance, but also its ability to hold disparate data together and discursively organize them into a single image that, in this case, refers to the subjects and subjectivities that first produced them.

It can be argued that since data images are already the result of an assumed impersonal as super-personal utterance, the semiotic operation of such an effect can be found in the narrative organization of information within such a type of images. In this context, I will look at one type of journalistic data image and try to understand the ways in which the subjectivity of the speaker—which would contaminate the authenticity of the communicated truth—is hidden and distorted by the discursive manipulation and the subjectivity of the people that is instead explicitly and rhetorically represented in the image.

In this respect, the data image appears as only one of the possible realizations of the virtual database to which the displayed information belongs.

² Within the same article Manovich argued that the narrative is just one method of accessing data among others.

4. Subjectivity in data images

In light of the preceding considerations, the notion that data images, as impersonal discourses, narratively produce an effect of truthfulness may seem novel. This is obvious in that, for semiotics, any discourse can be deconstructed according to a narrative logic. However, the narration of the data image is not only the precondition for the enunciative context of the visual discourse of the data image. Rather, it is about the narration strategies of modulating subjectivity in data images undertaken by empirical utterers.

The data images I will consider in the following chapter represent an absence, and by performing such an operation, they create a concatenation of utterances in which one point of view (that of the empirical utterer collaborating with the computational one) modulates another by assuming the freedom that is taken away from him/her (as in free indirect speech, Paolucci 2020: 254).

In this way, the subjectivity of the data referents is determined by a realization of the virtual information contained in the reference database.

The following data image is the result of a journalistic research conducted by *La Stampa* newspaper from Turin, Italy. In 2016, the “Torino Multe” project led to the visualization of over 4 million fines issued in the city between 2011 and 2015. For each of these fines, space-time information about the charges, the type of vehicle and the type of violation was organized in a database and an interactive visualization was created on the *Infractions* platform. Once the realization of the interactive platform was complete, the journalistic research work began.



Figure 1. Screenshot of the *Infractions* platform (www.datainterfaces.org/projects/infrazioni/)

As can be read on the website,³ the application consists of an API connected to the database of records, which is dynamically queried through the front-end interface that controls the display of the data. For example, when selecting the period between 2011 and 2012, it was observed that fines for driving under the influence of drugs were issued at checkpoints well known to citizens, especially in areas where nightclubs were concentrated; in contrast, in the following years, this type of violation was concentrated mainly in the area of San Salvario, a multi-ethnic neighbourhood of the city that began to fill with nightclubs since the early 10s. Another case concerned the fans of the Juventus stadium: from September 8, 2011, the day of the inauguration, until November 2012, no fines were imposed near the stadium; from that month on, fines suddenly accumulated. It results that was interpreted by the journalists of the Turin newspaper as a sign of a political decision in favour of a certain activity.

First of all, it should be noted that the fines do not refer to individuals, but rather to actions that are realized. They are impersonal in the sense that they are a collection, and the journalistic process can be interpreted in terms of the concatenation of utterances that contemporary semiotics speaks of. Such an operation can be described as highly deductive, since the data image merely offers a panoptic view of a series of events that have taken place in the area. It is precisely such an

operation that has produced an effect of *transparency* (ideologically understood) through the establishment of a distanced scopic observation regime.

If we consider such a case of journalistic data-visualisation discourse, we understand that the effect of truthfulness is narratively generated when the image succeeds in showing not only a mere visual representation of a condition (a referent); it succeeds when the image itself becomes a useful tool to infer such a referent from what is depicted.

A very different example concerns the *misleading* effect of truthfulness, presented by Alberto Cairo in his book *How charts lie* (2019).

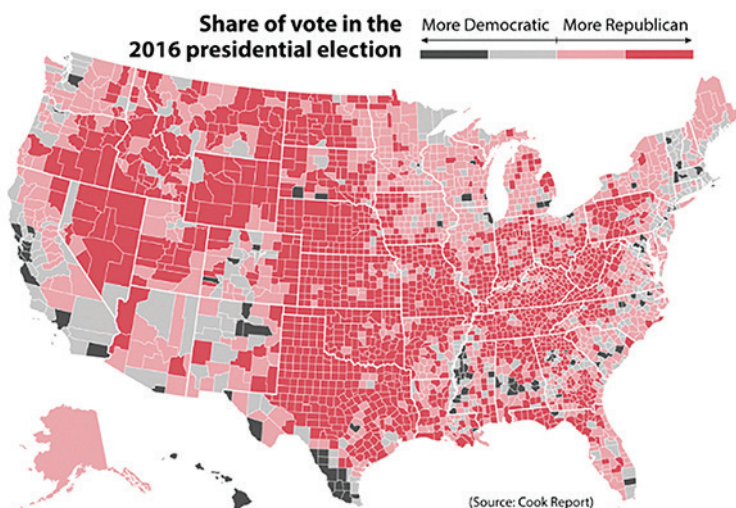


Figure 2. 2016 Trump's electoral map (by Cairo 2019)

³ www.datainterfaces.org/2016/07/infrazioni/

In analysing the journalistic disinformation strategies employed during Donald Trump’s run for the White House, Cairo found that the misleading images depicting Trump’s success did not actually represent the citizens who voted for that particular candidate, but rather it was the territories are represented.

Critics of President Trump were quick to excoriate him for handing out the county-level map to visitors. Why count the square miles and ignore the fact that many countries that went for Trump (2626) are large in size but sparsely populated, while many of the countries where Clinton won (487) are small, urban, and densely populated?

On the one hand, in this data image, the data on population size are potentialized, that is, taken out of the discourse and placed in the power of meaning; on the other hand, the data on territorial size are made relevant and *actualized*. We are thus dealing with a semiotic distortion, that is, the actualization of a virtualized form while empowering a realised form (Paolucci 2020: 212-4). It could be argued, then, that such discourse depersonalised people and personalised territories by achieving in this way the distorted representation of a desired truth.

Firstly, such an operation betrays the super-personality of the computational utterance, and does so precisely by publishing a non-truth account. It thus becomes a Trump’s utterance and not just a computational utterance. Secondly, it fails to present a transparent panoptic vision of the electoral situation. It does not fulfil the image’s purpose of allowing users’ self-reflection into the image.

The two images analysed are examples of data journalistic imagery, a text typology prototypically used in a social domain highly dedicated to the discovery of truth. Although this is not actually a treatise on journalistic rhetoric, the hard perspective of the second image must be emphasised. Not only does the data picture fail to reflect reality, it distorted by representing only one discourse—among many—that in favour of Trump.

In this respect, I would argue that the effect of truthfulness could also be generated by representing a set of heterogeneous discourses capable of providing a more complex (and multi-perspectival, i.e., pluralistic) representation of reality.

A different path can be taken to describe the next two data images, which are atypical in the sense that, unlike the previous ones, they explicitly defend a particular ideology, but without aiming at the value of a scientific discourse, thus generating a truthfulness effect.

Moreover, like the first one, these images share an important characteristic: they are interactive. This characteristic does not prescribe a passive reading practice that could be compared to the scientific assertion of referential, but an active role for the reader that ultimately offers the possibility of personalising the viewing experience and modulating the subjectivities that most interest the reader.

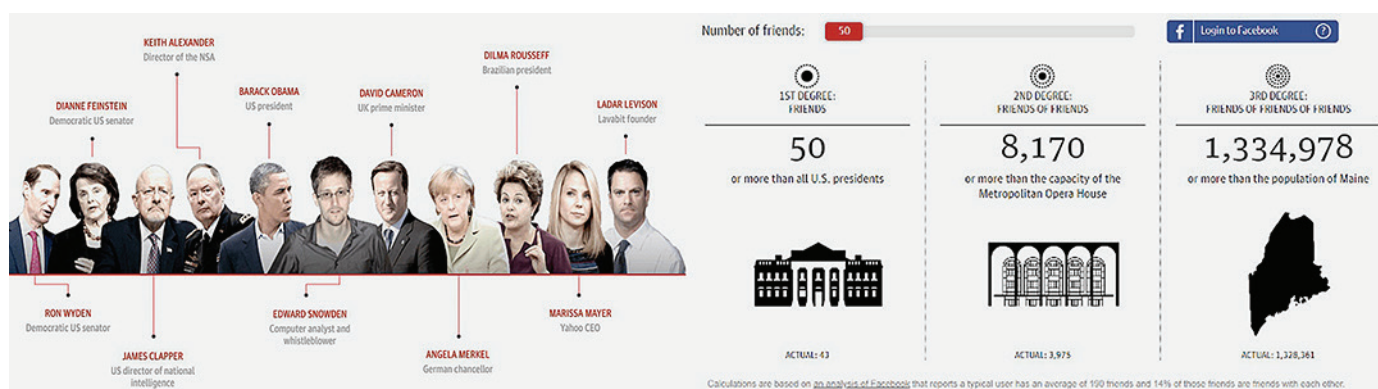


Figure 3. Journalistic data-images by The Guardian (www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded)

The two images again fall into the category of journalistic data visualization images: they were used by *The Guardian* to support the Snowden revelations about NSA surveillance programs in 2013.⁴ In the image on the left, you can see an interactive visualization of what some public figures said to the press following those revelations. By clicking on the various characters, you can read and hear what each person said by juxtaposing them in real time. Even if it is not a real numerical data image, it creates the effect of truthfulness in the representation of a plurality that, as a rhetorical consequence, assures the super-personality of the utterer (in this case *The Guardian*). Indeed, such an image is not merely the result of a computational representation, but the effect produced must be understood as a social value inscribed by the journalistic actor himself, concealing his own subjectivity in the subjectivities represented. This is a perfect example of the effect of truthfulness generated by the representation of heterogeneous subjectivities, which nevertheless affirms the super-personality of the utterer. Seen in this light, it can be said that journalistic storytelling has always produced the effect of truthfulness through the multiplication of points of view.

Although, what about the second image? It asks users to log on to Facebook to see how much personal information is retrieved by the computerized mass surveillance system. In this case, the computerized information retrieval is simply based on the number of friends one has on the social network's platform. In addition to the interpellation effect resulting from the call to action and the rhetoric of fear-mongering through the figurative comparison with ever larger objects, this is a good example of real-time visualization of personal data, such as the statistics automatically generated by SNS algorithms about the actions performed by users in a given period of time.

⁴ <https://www.theguardian.com/world/interactive/2013/nov/01/snowden-nsa-files-surveillance-revelations-decoded>

Finally, a type of data image that can be said to drastically remove human subjectivity from its referent is represented by a very suggestive image presented by Google on Earth Day as one of its doodles. According to its website, “the doodle uses real time-lapse images from Google Earth time-lapse and other sources to show the effects of climate change in four different locations on our planet; stay tuned throughout the day to see these scenes, each of which will be on the homepage for several hours.” In particular, the image below shows the Harz forests destroyed by bark beetle infestation due to rising temperatures and severe drought.

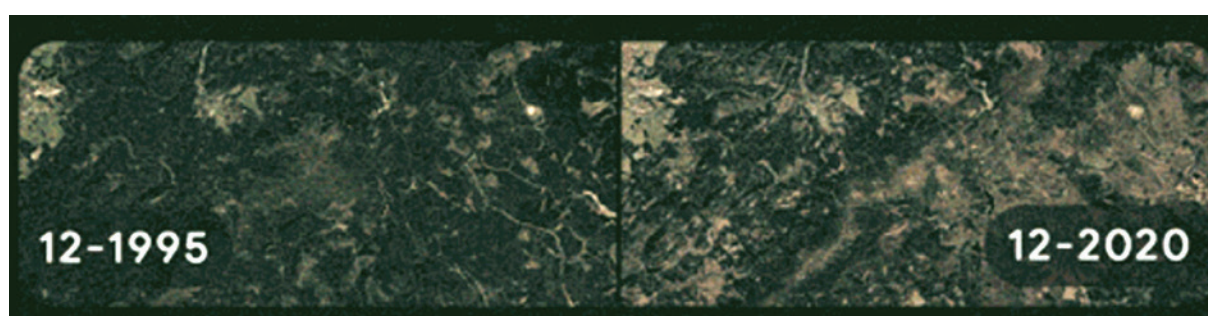


Figure 4. Google Earth Day’s doodle

In line with the previous analysis, it can be argued that this image does not modulate subjects, but rather *hyper-objects* (Morton 2015). However, such a claim might be considered unimportant in comparison to the visual rhetoric it implements—which the enunciative semiotic analysis of the previous images would have considered marginal.

This image clearly creates a narrative effect of emergence given by the *increasing plastic linearity*. The user can move the cursor over the horizontal plane to observe the gradual iconic transformation expressed in a non-neutral visual effect that highlights the excess, the density compared to more sober years—the 1990s—and can touch the imagination of the users using Google that day. In this case, it is not about the information of the data, but about its *substance*, its density and its respect for an old and nostalgic, almost romantic sobriety.

This is not the context in which we are dealing with scientific images: nor is it a scientific image, but a very rhetorical one with a precise discursive goal. It is ultimately an iconic representation of the thematic isotopy of data pollution. Although there are very few examples of discourses that focus strictly on “data pollution”, such a concept refers mainly to the pollution that results from “information overload”—in this context, we often speak of information *glut* (Shenk 2013) and information *diet* (Johnson 2012), referring not so much to the media as an organism, but to the living, biologically determined being and its embodied relationship to the device, which is nothing more than a window on the world, a heterotopic space that can also pose ecological problems.

5. Conclusions

This article aimed to be an exercise in applied semiotics, underpinning the various ways in which subjectivity can be modulated in data images.

However, it also opens up a perspective more devoted to the philosophical aspects. The hypothesis of a “realised” encyclopaedia—a container of digital traces and crumbs—reveals not only a set of socio-political problems, but also the risk which the discourses of contemporary digital culture increasingly express at the pragmatic level of computational enunciation.

On a meta-semantic level, Eco’s encyclopaedia itself—the paradoxical container of simultaneously virtual and real norms, customs, and patterns—becomes the object of a whole series of discourses and narratives (again, socio-political, but not only, also philosophical) that—if they are not only subversive (as in, when they are not merely subversive (as, for example, in the critique of gender stereotypes of language) take the form of attempts at collective consciousness-raising that are not so much interested in revealing the cultural determinations of the subject’s utterance as they are in the causes and consequences of a hypothetical “modulation mania”—a further development of the storage mania mentioned by Manovich—of which virtual reality is an explicit expression.

References

- Berry, D. 2011. *The philosophy of software*. New York: Palgrave Macmillan.
- Bihanic, D. (ed.), 2015. *New challenge for data design*. London: Springer-Verlag.
- Cairo, A. 2019. *How charts lie*. New York: Norton & Company.
- Chun, W. H. K., Cramer, F., Steyerl, H. and Apprigh, C. (2018). *Pattern discrimination*. Minneapolis: University of Minnesota Press.
- Cramer, F. 2005. *Words made flesh*. Rotterdam: Piet Zwart Institute.
- Dondero, M. G. 2020. *Les langages de l’image. De la peinture aux Big Visual Data*. Paris: Hermann.
- Drucker, J. 2014. *Graphesis*. Cambridge: Harvard University Press.
- Eco, U. 1979. *Lector in fabula*. Milan: Bompiani
- Eco, U. 2007. *Dall’albero al labirinto*. Milan: Bompiani.
- Eugeni, R. 2021. *Capitale Algoritmico: cinque dispositivi post-mediali (più uno)*. Brescia: Editrice Morcelliana.
- Ferraris, M. 2021. *Documanità. Filosofia del mondo nuovo*. Rome-Bari: Laterza.
- Fontanille, J. 1994. De simulacres de l’énonciation à la praxis énonciative. *Semiotica* 99: 273-291.
- Greimas, A. J. 1984. *Du sens: Essais sémiotiques, t. 2*. Paris: Seuil.

- Greimas, A. J. and J. Courtés 1979. *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
- Grusin, R. 2018. Datamediation. In: Montani, P., Cecchi, D. & Feyles. M. (eds.), *Ambienti mediali*. Rome: Meltemi, 165-180.
- Johnson, C. 2012. *The Information Diet. A Case for Conscious Consumption*. Newton (MA): O'Reilly Media.
- Kitchin, R. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Thousand Oaks: SAGE Publications.
- Manovich, L. 2001. *The language of new media*. Cambridge, MA: MIT Press.
- Moretti, F. 2005. *La letteratura vista da lontano*. Rome: Einaudi.
- Morton, T. 2013. *Hyperobjects: Philosophy and Ecology After the End of the World*. Minneapolis: University of Minnesota Press.
- Paolucci, C. 2020. *Persona. Soggettività nel linguaggio e semiotica dell'enunciazione*. Bologna: Bompiani.
- Pasquinelli, M. 2019. Three Thousand Years of Algorithmic Rituals: The Emergence of AI from the Computation of Space. *E-Flux Journal* 101.
- Pentland, A. 2014. *Social Physics: How Social Networks can make us Smarter*. London: Penguin Books.
- Reyes, E. 2017. *The Image-Interface: Graphical supports for visual information*. Hoboken: Wiley.
- Severo, M. and Venturini, M. 2016. Intangible cultural heritage webs: Comparing national networks with digital methods. *New media & society* 18(8): 1616-1635.
- Shenk, D. 2013. *Data Smog: Surviving the Information Glut*. San Francisco: Harperone.
- Sieckenius de Souza, C. 2005. *The Semiotic Engineering of Human-Computer Interaction*. Cambridge, MA: The MIT Press.
- Willer, B. 2015. Show don't tell. In: Bihanic D. (op. cit.), 3-22.

AUTHOR

Federico Biggio PhD in Semiotics and Media, lecturer and researcher in SIC at Université Paris VIII.

