Data Visualization as Portraiture: An analysis of the use of personal data in the visual representation of identity

Catarina Sampaio
hello@catarinasampaio.com
Universidade de Lisboa,
Faculdade de Belas Artes,
Centro de Investigação
e de Estudos em Belas-Artes
(CIEBA), Lisboa, Portugal

Luísa Ribas
l.ribas@belasartes.ulisboa.pt
Universidade de Lisboa,
Faculdade de Belas Artes,
Centro de Investigação
e de Estudos em Belas-Artes
(CIEBA), Lisboa, Portugal

This paper analyses aesthetic and design artefacts that represent the identity of an individual through the visualization of personal data, focusing on their reflexive, documental, and biographical potential. To this end, it analyses the concepts, data types, design methodologies, and visualization techniques, as well as the experiential characteristics of these artefacts. This analysis is part of an ongoing research that explores data visualization techniques for the representation of personal identity, highlighting the potential of using personal data as a raw material for portraiture. The research is motivated by the current context of technological ubiquity, wherein virtually all human activities inevitably leave a digital trace. It explores the potential of these records of personal data to convey characteristics of identity that are relevant in private and social spheres. As such, it also points to a reconceptualization of portraiture driven by computational media, as the representation of personal identity through the use of personal data.

1. Introduction

Our current daily activities are constantly mediated by a growing number of technological devices capable of quantifying and recording various aspects of human life in the form of digital data. These records can include vast sets of highly personal information about actions, habits and preferences, social and intimate interactions, as well as bodily functions. These data refer to specific individuals who can be identified through them (Wolf 2010). Drawing on the notion that "new media offer new opportunities for self-definition" (Bolter and Grusin 2000, 231) this research explores how the visualization of personal data can efficiently convey characteristics of one's identity.

In doing so, it also seeks to highlight the wider uses and implications of personal data, by addressing issues related to one's agency and control over personal data (Kennedy et al. 2015; Mun et al 2010). In this sense, authors such as Stepanchuk (2017) point to a growing hegemony of entities that facilitate the development of technology that enables the collection and archiving of personal data. In this context, they highlight how the construction of identity becomes conditioned by the affordances of technology, as a result of technological determinism (Poletti and Rak 2014). They point to the implications and, not yet fully predictable effects, that digital technologies have on human ontologies (Hernández-Ramirez 2017; Floridi 2014), as they structure an individual's relationship with themselves and with others.

Authors like Zuboff (2019) and Harary (2015) highlight issues of commercial exploitation and social manipulation through the collection and analysis of personal data. But despite the current context of surveillance capitalism highlighted by these authors, there are alternative views and creative approaches to the use data produced by and about individuals. Designers like Lupi (2017) explore data from a humanist perspective, through the creation of visualizations that support decision-making processes (Ciuccarelli et al. 2014). These approaches also seek to facilitate social interactions between users in digital environments (Donath 2017) or seek to provide the individuals with mechanisms of self-knowledge that allow them to act upon themselves (Foucault 1998; Floridi 2011). Motivated by these ideas, this paper is part of an ongoing research that explores the concept of "data portrait" as coined by Xiong and Donath to describe "representations of people made by visualizing data by and about them" (Donath 2017, 187).

2. Data Portraits

The first notable experiences around the concept of data portrait emerged in the late 1990s, to represent participants in online forums. In this context, visualizations were created by mapping each user's activity to help group members

1. The *PeopleGarden* (1999) project was one of the first data portraits created at MIT Media Lab by Rebecca Xiong and Judith Donath. It used the "simplicity and intuitiveness" of the garden metaphor to "convey how user's behaviour changed over time" (Xiong and Donath 1999).

2. For example, *Portrait of Deb from 1988-199?* (2012-2013) by L. J. Roberts, results from a series of embroidered emblems collected by Deb, as the outcome of an autoethnographic practice. The work is based on the notion that material objects express meaning and employs enumeration techniques and personal inventory as a form of portraiture (Goodyer and al. 2016, 240). Another example would be *Everyone I Have Ever Slept With 1963-1995* (1995), by Tracey Emin, which is an art installation composed of a tent with the appliquéd names of, literally, everyone she had ever slept with, including family, friends, lovers and even two numbered foetuses (West 2004, 211).

"to make sense of each other" (Donath 2017, 198).[1] A data portrait can then be defined as a graphic representation of personal identity achieved through the visualization of personal data. It is the result of a multidisciplinary practice that essentially applies data visualization techniques to the realm of portraiture. This approach to portraiture is reminiscent to forms of artistic production that appropriate techniques from scientific disciplines, such as the exhaustive observation of daily life, data collection, analysis, and inventory, as an ethnographic basis to produce aesthetic artefacts (Morley 2007).[2]

Addressing the notion of data portrait implies debating the relationship between *portraiture, data* and *visualization* (Sampaio et al. 2019). Commonly, a *portrait* seeks to evoke an individual by visually conveying characteristics of their identity (West 2004). This can be done by depicting similarities related to their physical appearance. It can also be accomplished by resorting to contextual elements, such as certain poses, clothing, and everyday objects that illustrate their habits and interests, and thus convey the identity of the person they represent (West 2004). For Xiong and Donath (1999), the online equivalent of these contextual elements is behavioural data. *Data* are "abstract in nature and in themselves are meaningless. Only when organized and contextualized do they produce information" (Lee 2014, 19). Thus, data can be considered raw material from which information emerges, through cognitive processes and visual perception (Meirelles 2013). Accordingly, *visualization* can be understood, on the one hand, as the formation of ideas resulting from the human ability to reason, that is, to the creation of a mental image. On the other hand, it refers to the materialization of an image that allows to represent something visually, that is, the conversion of concepts into images or visible forms (Almeida 2017). In the field of information design, visualization is a representation technique that comprises the mapping of abstract data into a graphic system to produce meaning (Manovich 2002).

As we have discussed elsewhere (Sampaio et al. 2019), the visualizations of personal data addressed in this paper can be understood as 'portraits' because they visually communicate personal identity by fulfilling biographical and documentary functions. Nevertheless, they differ from most traditional forms of portraiture because, rather than conveying identity through physical appearance (conveying gender, age, or race), they do so by highlighting behavioural traits and patterns related to one's lived experience, habits and interests. In this sense, data portraits do not describe one's *appearance* but evoke one's *experience.*

3. Functions of data portraits

Data portraits take advantage of the potential of personal data as the product of one's daily experiences to convey identity traits. However, rather than merely assuming an analytic stance, data portraits also value the subjective expression

of what is commonly understood as objective, numerical data as an index of reality.[3] In this sense, they resort to "techniques from the universe of statistical analysis, but their purpose is artistic" (Donath 2017, 209).[4]

As designed artefacts, data portraits also fulfil specific expressive and communicative functions that are similar to their traditional counterparts, as an "artistic" endeavour with a "biographical" or "documental" function, as a "proxy", or even, as a "political tool" (West 2004, 43-69).[5] In this sense, data portraits can act as a *proxy, double or substitute* for users in virtual environments (Donath et al. 2014). They can also work as a *mirror or a self-reflection* that allows the observation of behavioural patterns (Sampaio and Ribas 2021). Additionally, they can be a *critical discourse* and promote awareness of the wider implications of our digital footprint and our loss of control of the data produced by and about us (Donath et al. 2014).

3.1. The data portrait as double

To establish interpersonal relationships in virtual environments we need interfaces that allow for self-representation, because personal identity is not embodied. In this context, the *selfie* has become the preferred way of "bringing the face to the interface" (Donath 2001). However, despite being a mimetic representation, its use is more advantageous in social networks where users know each other in real life, because, without this reference, the *selfie* loses its contextual value and can become ambiguous or even misleading (Donath 2001). Consequently, alternative modes of self-representation based on quantifying behavioural data, such as data portraits, can be less prone to be elusive or misinterpreted (Rettberg 2014).

Xiong and Donath (1999) identified several advantages in representing individuals in online communities through their behavioural data. One of them is that this kind of representation offers context, comprising a history of past interactions within the community. Also, user status can be updated according to their most recent interactions. As such, these visualizations can evolve and change over time, sometimes in real-time. Being based on objective and non-anonymized data these portraits can be considered reliable. At the same time, they also provide some degree of privacy, as they produce meaning only within the context of the platform on which they are generated (Donath et al. al. 2014). Finally, as they do not rely on mimetic attributes, they ideally eliminate stereotypes since assumptions about the subject/user are based on their real actions and not on potentially biased characteristics such as sex or race (Donath 2001).

3.2 The data portrait as a mirror

In addition to conveying socially relevant aspects of identity, and due to their biographical and documental nature, data portraits can also fulfil a reflexive function by expressing traits of identity that are most significant within the personal sphere. In this case, the data portrait is designed to act as a "data mirror" reflecting one's behavioral patterns (Donath et al. 2014).

Much in the same way several artists have used the self-portrait as an exploration of their identity throughout history (West 2004), modern self-tracking technologies also facilitate this type of biographic record through data collection and visualization. When coupled with self-tracking techniques, the data portrait can materialize as an interface between the user and their data. Like a traditional self-portrait, it can be a means for self-knowledge and exploration. In this sense, data portraits can be framed as "technologies of the Self" (Foucault 1998). As Foucault argues, these technologies allow individuals to "establish a closer and truer relationship with themselves" (1998, 18). They belong to a set of specific techniques that human beings use to produce knowledge about themselves and to implement changes on their own lives in order to optimize them (Lupton 2014).

According to Donath et al. (2014), in the private sphere, data portraits function as "data mirrors, portraits designed to be seen only by the subject, as a tool for self-understanding". This reflexive function results from the affective relationship that the user creates with data when visualized (Pink et al. 2018). According to studies carried out by Lupton (2016)[6] and Pink et al. (2018), personal data, when visually represented, acquires a symbolic value and generates feelings of belonging, which are associated with memories and a notion of continuity. As with photographs displayed in an album, data portraits thus acquire a sentimental value due to their biographical and documental qualities.

3.3 The data portrait as critical discourse

In most of our daily activities, in addition to the information we intentionally create and give away, vast records of data are generated without our full awareness, resulting in a digital footprint that is also accessible to other parties, such as government agencies, commercial companies, and data brokers (Rettberg 2014). However, we have little or no access to those data or the representations made of us, even though they are the product of our actions and interactions (Harary 2015).

Feeding on our "behavioural surplus" (Zuboff 2019)[7] algorithms are constantly analysing our identity, by translating our actions, preferences, and interests into behavioural data. This process of data extraction began online, resorting to metadata as "device/software generated data" that is "necessary for every activity on the internet" (Joler et al. 2015a), but is continually being driven towards

5. West (2004, 43-69) summarizes the functions of traditional portraiture as "the portrait as a work of art", "the portrait as biography", "the portrait as document", "the portrait as proxy or gift", "the portrait as commemoration or memorial" and, finally, "the portrait as political tool". As proposed by West, portraiture is traditionally understood as serving "functions that other works of art have not (…) in a way that is unique to its genre" (West 2004, 220) given that portraits "are representations, but also material objects" (West 2004, 43).

6. Lupton expands on the idea that the materialization of personal data through its visual representation connects data with users on her book Quantified Self (2016).

7. Behavioural surplus is described by Zuboff (2019) as data that goes beyond online product and service use.

new sources in the real world. Metadata is indispensable for communication on the internet as we know it today, but it is also extraordinarily intrusive. As stated by Edward Snowden in a conference call in 2014, "Metadata absolutely tells you everything about somebody's life. If you have enough metadata, you don't really need content" (Joler et al. 2015b). According to Zuboff, in the context of "surveillance capitalism", the use of these data goes far beyond online product and service improvement, as it is now being primarily claimed "as free raw material for hidden commercial practices of extraction, prediction, and sales" (Zuboff 2019, 8).

In this context, data portraits may provide a better balance between the intentional ways we convey our identity and the unintentional traits of identity we provide to third parties. By making our digital footprint visible and accessible, data portraits can be a means to attain some degree of agency over personal data making it "accessible, understandable, and actionable for our everyday tasks" and, as such, connect "people with timely and meaningful insights" (Bonde 2013).

Data portraits can also build a critical discourse on the established power relations around personal data, evoking current debates on data protection and privacy, and the common disregard of big data companies of user's rights and control over private information. As such, data portraits also fulfil a political function when devised as critical statements on current data commodification practices and data policies. According to this view, data portraits become not only an expression and manifestation of an individual's presence in the world but also "a point of view on that world and a way of potentially recreating or restoring it", considering that, as stated by Medeiros (2000, 36), "to represent is always to revolutionize".

4. Analyses of data portraits

In order to further understand the expressive and communicative functions of data portraits, as a double, self-reflection and according to their critical stance, we analysed a set of artifacts concerning their different creative approaches to the visual representations of identity based on personal data. The current analysis updates and expands a previous one (Sampaio and Ribas 2020) by delimiting the scope to artefacts that use computational means for collecting, processing, and visually representing data, to highlight the mediation of identity by technologies of everyday use.[8]

4.1 Scope

The current selection of aesthetic and design artefacts corresponds to works that 1) aim at the visual representation of personal identity, 2) resort to the visualization of personal data, 3) are based on computational media and 4) are not

8. As sources of information on the projects, we resort to direct observation, authors' documentation, and bibliographic references to the analysed works, as in the previous study (Sampaio and Ribas 2020).

visually mimetic. This selection seeks to comprise a diversity of projects, ranging from the emergence of the concept of data portrait, as established by Xiong and Donath in 1999, to the present day:

1. *PeopleGarden*, Rebecca Xiong & Judith Donath, 1999
2. *Authorlines*, Fernanda Viégas & Marc Smith, 2004
3. *A Week in the Life,* Andreas Fischer, 2005
4. *The Dumpster*, Golan Levin, Kamal Nigam and Jonathan Feinberg, 2006
5. *Themail*, Fernanda Viégas, Scott Golder and Judith Donath, 2006
6. *Fitbit*, Gadi Amit et al., 2007 – ongoing
7. *Personas*, Aaron Zinman and Judith Donath, 2008
8. *Spigot (Babbling Self-Portrait)*, Jason Salavon, 2010
9. *Atlas of the Habitual*, Tim Clark, 2010-2011
10. *TimeMachine*, CADA, 2012
11. *Nike Hyperfeel Experiment,* Aramique, 2013
12. *Heart Bot*, Aramique, 2014
13. *Data as Object*, Brendan Dawes, 2014
14. *The Art of the Thrill*, Soso Limited, 2014
15. *Porsche BlackBox*, OnFormative, 2015
16. *Good Night SMS*, Paul Heinicker, 2015
17. *The Sixth Sense*, Clever Franke, 2016
18. *Poisonous Antidote*, Mark Farid, 2016
19. *Floating Map (Life Location Project)*, Stephan Cartwright, 2016
20. *Heart of Travel*, Joshua Davis, 2017
21. *Data Selfie*, Data X, 2017
22. *Halo*, Peter Crnokrak – ORA, 2017
23. *The Art of Feeling*, Random Quark, 2017
24. *Building Hopes*, Accurate, 2018
25. *OnePlus 7 Live wallpaper*, OnFormative, 2019
26. *Made to Measure*, Studio NAND, 2020

4.2 Aim and focus

Considering the main communicative and expressive functions of data portraits, the framework for analysis focuses on 1) the work's *concept*, i.e., its motivations, objectives, and approaches to personal data, 2) on the work's *mechanics*, i.e., data collection and analysis processes in order to analyse how the work's intent is actually implemented, and 3) on the elements of their *experience*, i.e., the outputs, interfaces and behaviour of the work, addressing the main characteristics of the visual representation and the kind of perception of the source data these portraits allow.[9]

9. The analysis framework was adapted and expanded from a previous study (Sampaio and Ribas 2020) and based on a model proposed by Ribas (2014) for the analysis of digital computational systems. It is also based on the model that describes the process of deriving information from data through its visualization described by Fry (2008), which in turn systematizes the "informational-digital praxis" proposed by Renaud (2003).

In particular, we consider how the choice of personal data defines the *theme* and *content* of the data portrait concerning their functions. We then observe the *data collection* and *data analysis* processes, which are implemented at the level of their mechanics, as specific means to accomplish the work's intent. Finally, we observe the elements of the *experience* of each work, or what is accessible for the user/audience to experience, in terms of *surface* (outputs and interfaces) and *dynamics* (behaviour), which define the perception and experience these visualizations promote (Table 1).

Considering these specific points of view on the works analysed, we now describe the categories of the framework, as informed by parameters referenced below and as resulting from a comparative analysis of the works selected.

Concerning the works' *concept* and theme or subject matter, the focal point of the current analysis is the portrait's intent as tied to the main functions of the data portraits: as a *double* (a proxy of the user in online communities), or as a data *mirror* (a mechanism of self-observation through self-tracking), or as a *critical discourse* (by re-appropriating personal data). Consequently, we find it useful to specify different *types of portraits: self-portraits, individual portraits, and collective portraits.* We also examine the *motivation* behind the creation of the portrait, as an *internal* motivation tied to a personal discourse or artistic reflection, or whether it results from an *external* motivation, such as with the development of a product or service.

Within the conceptual dimension, we also consider the works' *content*, as tied to its function and relating to the *data sources* used, whether collected from a *virtual environment* or the *real world*. The *typology* of data is described in terms of *behavioural data, biometric data, data related to the surrounding environment.* A relevant aspect to this analysis is the *ethnographic value* of the data used in the visualization that, according to Whitehead (2005), can express a wide array of information about the individual portrayed, namely their *physical characteristics, psychological traits, social relationships, behavioural patterns, or their physical environment.*

Table 1. Analysis framework.

1. CONCEPTS		2. MECHANICS		3. EXPERIENCE	
THEME	**Portrait's intent** Data double Data mirror Critical discourse **Type of portrait** Self-portrait Individual portrait Collective portrait **Motivation** Internal External	**DATA COLLECTION**	**Data collection** Service providers Devices Web 2.0 **Level of control** Deliberate Non-deliberate **Recording format** Textual Numerical Other	**SURFACE**	**Informational structure** Hierarchical Relational Temporal Spatial Spatio-temporal Textual **Representational stance** Analytical Expressive **Modes of expression** Visual Audiovisual Material **Output format** Static image Physical object Moving image Digital interface
CONTENT	**Data source** Virtual environment Real world **Typology** Behavioural data Biometric data Surrounding environment data **Ethnographic value** Physical characteristics Psychological traits Social relationships Behavioural patterns Physical environment	**ANALYSIS**	**Data analysis** Statistical None Not possible to observe	**DYNAMICS**	**Behaviour of the output** Contemplative Interactive **Nature of the output** Static Transient Variable

Moving on to the *mechanics* of the work, or the specific implementation of its conceptual aim, we examine how the *data collection* is done. Namely, through *service providers* using digital network services (big data), *devices* with built-in sensors (IoT), or provided by the user in *social networks* and other similar platforms (web 2.0). A relevant aspect to the current analysis is observing the *level of control* that the subject/user has over data collection processes. Based on Selke (2016), we distinguish between a *deliberate* way of collecting data (i.e., using self-tracking tools) and a *non-deliberate way* (when data collection is automated without the active intervention of the user). The data *recording format* is also considered, if it is *textual, numerical,* or *other* (i.e., image, sound), as well as the processes of *data analysis* implied, to highlight if *statistical analysis* or other methods are applied to personal data.

Finally, considering the *experience* of the data portrait, we focus on the sensory expression and observable behaviour of the work, to understand the kind of perception and aesthetic experience it promotes. We observe the type of *informational structure* employed in the visualization according to the categories defined by Meirelles (2013), which can be *hierarchical, relational, temporal, spatial, spatio-temporal,* or *textual*. Additionally, the representational *stance* of the visualization can be *analytical* (if the visualization favours legibility) or *expressive,* when it

favours a subjective experience of the data, or an aesthetic experience promoting reflection and inciting emotional response. The *modes of expression* of the output can either be *visual, audio-visual,* or *material*, and the *output format* of the visualization system can be a *static image*, a *physical object*, a *moving image*, or a *digital interface*. Additionally, the nature of the *output* can be either *static, transient* (time-based) or *variable* (real-time). Finally, regarding the *dynamics* of the data portrait, the overall *behaviour* of its *output* can be merely *contemplative* or devised as an interface allowing *interactive* exploration.

5. Tying concept, implementation, and experience

The results of this analysis point to the interdependency between conceptual intent, relating to the functions of data portraits, and the types of data and visualization strategies used, within a diversity of modes of expression and formal aspects of data portraits.

5.1 Approach to personal data

We observed that these artefacts often articulate several expressive functions, with a predominance of the reflexive one. Their theme is always shaped by the personal data that visualizations use as content, which also relates to their typology (Table 2). Data portraits of the late 1990s and early 2000s were predominantly aimed at representing users in online communities and were, at the same time, individual and collective portraits (e.g., *PeopleGarden, Authorlines, Themail*) whose goal was to act as a double or proxy for the subject/user, communicating their identity to others. Subsequent data portraits are more diversified thematically, also articulating various expressive functions.

Table 2. Analysis results — concepts.

Portrait's intent	Type of portrait	Motivation	Data source	Typology	Ethnographic value
Data double	Self-portrait Collective portrait	External	Virtual environment	Behavioural data	Social relationships Behavioural patterns
Data mirror	Self-portrait	Internal	Virtual environment Real world	Behavioural data Biometric data Surrounding environment data	Physical characteristics Psychological traits Behavioural patterns Physical environment
Critical discourse	Self-portrait Individual portrait	Internal External	Virtual environment	Behavioural data	Physical characteristics Psychological traits Social relationships Behavioural patterns Physical environment

Many of the projects analysed aim to represent individuals before themselves, as self-portraits or self-observation tools through self-tracking (e.g., *Fitbit, Time-Machine, Halo*). Some data portraits depart from this function to then construct

a critical discourse (e.g., *Spigot, Data as Object, Poisonous Antidote*). These portraits fulfill a political role, based on the aggregation and re-appropriation of personal data that is scattered on different networks and platforms. Such as the case of *Spigot (Babbling Self-Portrait)* or *Poisonous Antidote,* which highlight aspects related to lack of privacy and security in the online world, raising questions about who owns these data. Projects with this type of critical stance tend to explore issues related to the individual's agency over their data and are usually moved by the internal motivation of their creators. In turn, projects that primarily aim at conveying personal identity, fulfilling the function of data double, often stem from external motivations and therefore have a more applied nature.

Regarding the data sources related to the portraits' intent, we observed that when the data portrait fulfils the role of a *double,* or a proxy communicating the subject's identity to others, it tends to be based on behavioural data retrieved from user activities within a platform. Such is the case of *PeopleGarden* and *Authorlines* that resort to the visualization of these types of data to hinder the deceitful self-representation of users in these online communities (Donath 2017). On the other hand, when the portrait is self-directed and fulfils a *reflexive* role, it usually resorts to biometric data, sometimes correlated to data of the individual's surrounding environment, as data extracted from the real world.

Most data portraits allow us to infer patterns of behaviour, and sometimes couple these with psychological traits, social relationships, or the environment of the individual portrayed. When biometric data is correlated to data of the surrounding physical environment, it allows us to infer the psychological traits of the subject. Such is the case of *The Art of Feeling,* which uses EEG signals to detect positive or negative emotions, and The *Art of the Thrill* that visualizes the excitement one feels when driving a sports car by measuring one's heart rate,

breathing, perspiration and body temperature. Finally, the more types of personal data are correlated in the visualization system, the more information it is possible to infer about the person portrayed, and the more ethnographically relevant the portrait becomes (e.g., *The Sixth Sense*).

5.2 Data collection and analysis processes

The different data portraits analysed involve diverse means of data collection and ways of processing data in the preliminary stages of their development, which determine their final outputs (Table 3). Biometric data and data on the surrounding environment are always captured by sensors embedded in self-tracking devices. If the portrait resorts to behavioural data it tends to use software (like web 2.0 services or platforms) or hardware, like sensors integrated IoT devices (e.g., smartphones or wearables like smartwatches, pedometers, and heart rate monitors).

Portrait's intent	Data collection	Level of control	Recording format	Data analysis
Data double	Service providers Web 2.0	Non-deliberate	Textual Numerical	Statistical
Data mirror	Service providers Devices	Deliberate	Textual Numerical	Statistical None
Critical discourse	Service providers Devices Web 2.0	Deliberate Non-deliberate	Textual Numerical Other	Statistical None

Consequently, when the data portrait is intended to act as a *double* or a proxy for the subject/user, the behavioural data are usually captured in a passive and non-deliberate way, meaning that the user does not have to actively interfere with the data collection process (e.g., *PeopleGarden*). On the other hand, when the portrait aims to fulfil *reflexive* functions and is self-directed, the user has some degree of agency in the process of data collection, using self-tracking devices and applications (e.g., *Fitbit*).

When the visualization assumes a critical stance, it tends to combine deliberate and non-deliberate methods of data collection (e.g., *Poisonous Antidote*). In this case, personal data is often reappropriated from web 2.0 platforms (e.g., *Data as Object*) and used as a means of understanding what these platforms actually 'know' about us (e.g., *Data Selfie*). Such is the case of *Spigot (Babbling Self-Portrait)* that uses data from Jason Salavon's browser search history, which Google keeps in file and uses to build consumer profiles. In most cases, personal data is handled by the visualization system, mostly, in numeric format, as a quantification of human experience or the real world.[10]

10. After being extracted from the human experience or the real world it is converted to a scale of pre-established numerical values. This quantification process aims to simplify the treatment of data by the computational media. Nevertheless, some of the projects also use data in textual format as a complement.
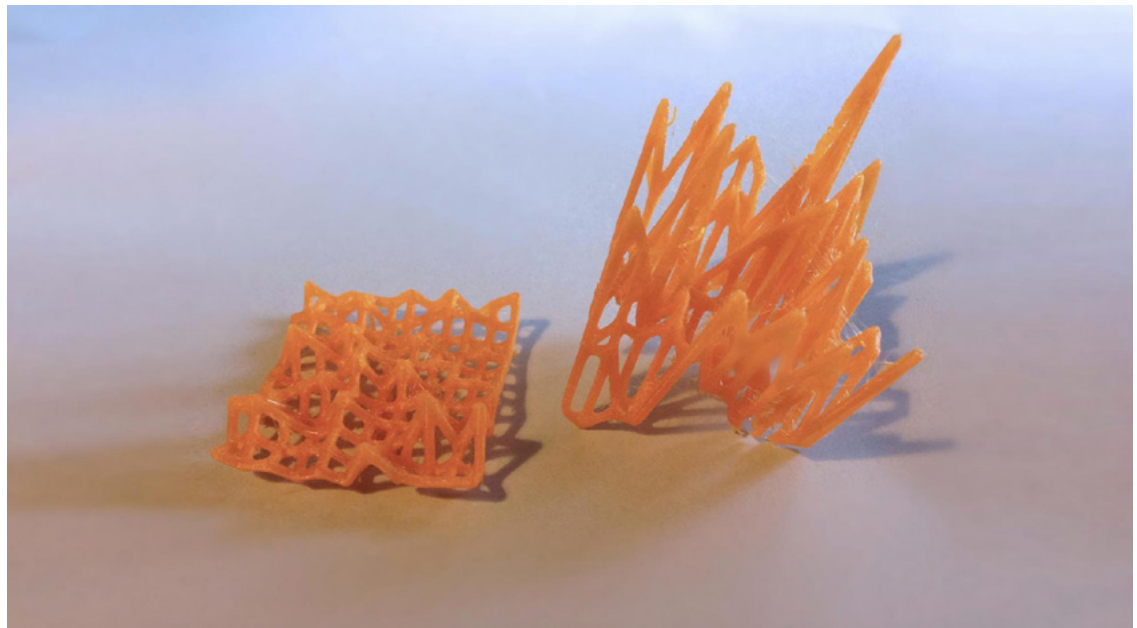
Consequently, and concerning the methods of data analysis, in addition to correlating various types of data to create ethnographically relevant data portraits, most projects also resort to the statistical analysis of data. They rely on quantified data to generate knowledge about the self, thus measuring various aspects of the individuals' daily experiences, and employing arithmetic operations to reveal behavioural patterns. This happens predominantly with portraits directed towards self-observation that resort to self-tracking (e.g., *Halo*), or in projects that seek to assess what information can be inferred by third parties from the data we dissipate on daily basis, even without our total awareness (e.g., *Data Selfie*). However, statistical analysis is not always involved, and projects like *Floating Map (Life Location Project)*, which translates geolocation data into geometric plexiglass sculptures, use personal data as raw material for creating expressive visualizations, emphasizing the creative possibility of mapping abstract data into any tangible form.[11]

11. These projects explore the visual representation of personal identity by employing the concept of transmutability of digital data that states that any kind of digital data can be translated into image or sound (Levin 2010, 273).

5.3 Visualization and expression of data

The diversity of modes of expression and formal aspects of data portraits reflect specific approaches to personal data, which can also be related to the portrait's function (Table 4). In this sense, most of the projects analysed use hierarchical, temporal, or spatio-temporal informational structures. Thus, the temporal dimension is often highlighted to express the variable nature of human experience over time and to favour the perception of patterns that emerge from data.

Table 4. Analysis results
— experience.

Portrait's intent	Informational structure	Representational stance	Modes of expression	Output format	Behaviour	Nature
Data double	Hierarchical Relational Temporal Textual	Analytical Expressive	Visual	Digital interface	Interactive	Variable
Data mirror	Relational Temporal Spatial Spatio-temporal	Analytical Expressive	Visual Audiovisual Material	Static image Physical object Moving image Digital interface	Contemplative Interactive	Static Transient Variable
Critical discourse	Hierarchical Relational Spatio-temporal	Analytical Expressive	Visual Audiovisual Material	Static image Physical object Moving image Digital interface	Contemplative Interactive	Static Transient Variable

However, despite using common information structures for data visualization, the representational stance often privileges expressiveness over legibility or analytical features (e.g., *Halo*). In contrast to visualizations of an analytical nature, data portraits value the subjective expression of data because, even if "the goal of visualization is often accuracy", the data portrait is also "an artistic production, shaped by tension among the often-conflicting goals of the subject, artist and audience" (Donath 2017, 187).

Given that the focus of the analysis is on visualizations (rather than sonification) the modes of expression are predominantly visual, be it moving images (e.g., *Porsche BlackBox*) or, sometimes, graphic interfaces (e.g., *Heart of Travel*). Complementary outputs, like prints (e.g., *The Sixth Sense*) or physical objects are also explored, in installation or performance settings (e.g., *Poisonous Antidote*), emphasizing the plasticity of digital data.

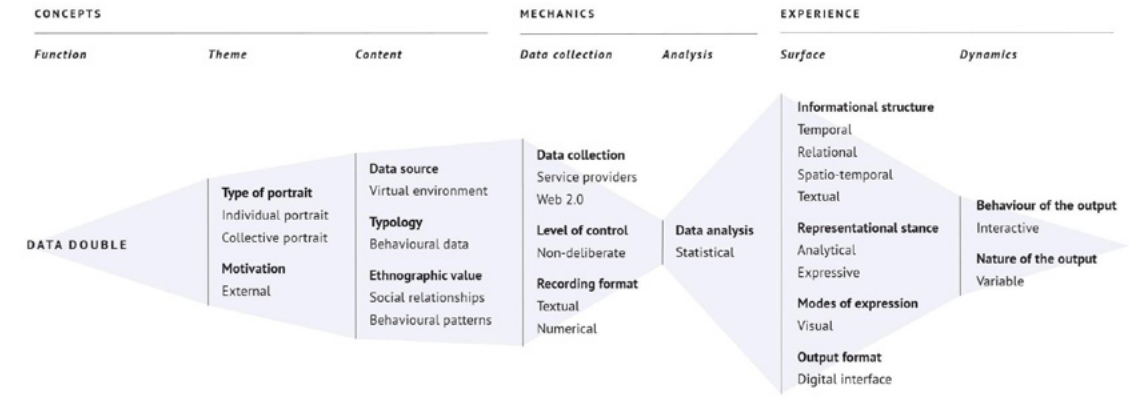Fig. 3. *Building Hopes*, Accurate, 2018.

According to these mostly static and transient outputs, data portraits are mainly contemplative, even though recent data portraits increasingly tend to have interactive properties, allowing users to explore the visualization at will (e.g., *Building Hopes*) or to manipulate its parameters (e.g., *Heart of Travel*). Thus, variable outputs are present across all kinds of data portraits, but they are still less frequent due to technical constraints since they require the system to respond in real-time to variations in the data flow.

6. Discussion

This analysis reveals how data portraits fulfil the same biographical, documentary, and reflexive functions inherent to traditional forms of portraiture. Although distanced from a mimetic stance based on physical appearance, they evoke the expressive and communicational functions of their classic counterparts as essentially tied to the representation of the subject before the other and/or before himself.

Data portraits that act as a *data double* or proxy of an individual in social contexts tend to have very clear and legible content and visualization strategies (Fig. 4). In this sense, the role of data double can be considered the most pragmatic function of data portraits, as applied to a specific context, which also helps to delimit the data they rely on, the data analysis, and the visualization strategies applied. Accordingly, it relies on behavioural data that is prone to reveal behavioural patterns, often collected through service providers and Web 2.0 platforms, in automated non-deliberate ways. Resorting to different types of informational structures used, their representational stance can nevertheless vary between analytical and expressive, and being applied to an online environment, the modes of expression are mostly visual, as interactive digital interfaces. The data portrait, as a double, aims at effectively communicating identity traits that are relevant to the social dynamics between users in virtual environments. As such, it is also a contextual portrait, both an individual and a collective portrait that emerges from the need of interfacing social interactions within online communities.
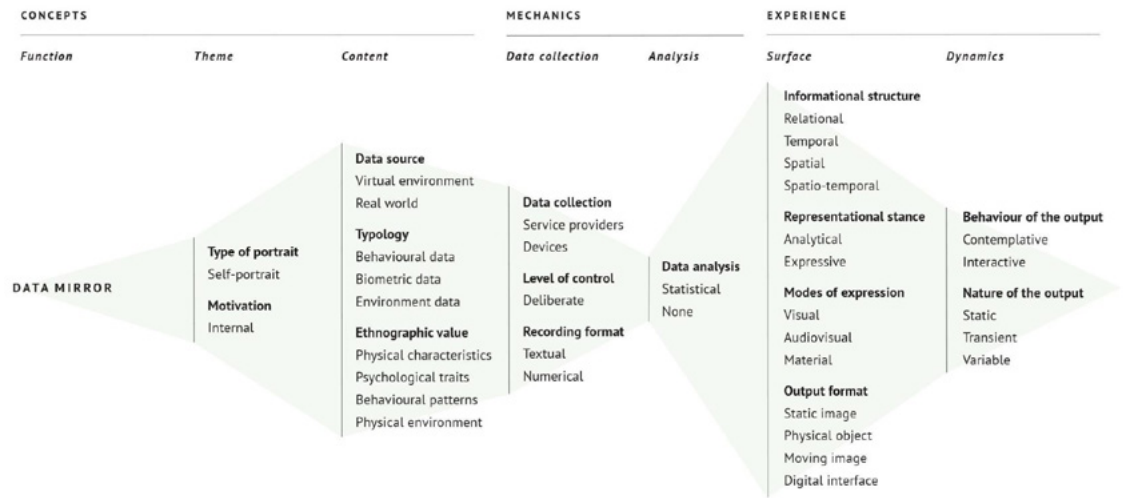
Fig. 4. Characteristics of the data portrait as data double.

CONCEPTS			MECHANICS		EXPERIENCE	
Function	*Theme*	*Content*	*Data collection*	*Analysis*	*Surface*	*Dynamics*

DATA DOUBLE

Type of portrait
Individual portrait
Collective portrait

Motivation
External

Data source
Virtual environment

Typology
Behavioural data

Ethnographic value
Social relationships
Behavioural patterns

Data collection
Service providers
Web 2.0

Level of control
Non-deliberate

Recording format
Textual
Numerical

Data analysis
Statistical

Informational structure
Temporal
Relational
Spatio-temporal
Textual

Representational stance
Analytical
Expressive

Modes of expression
Visual

Output format
Digital interface

Behaviour of the output
Interactive

Nature of the output
Variable

When acting as reflexive *data mirrors*, directed towards self-observation, data portraits convey relevant characteristics of identity that are valued mostly within the individual's private sphere (Fig. 5). As such, their themes are well circumscribed, as self-portraits that emerge from internal motivations. But the data that they resort to can be broader in scope in relation to the other functions of data portraits, comprising the virtual environment, and the real world, as behavioural, biometric, or environment data. The ethnographic value of the portrait is also broader since these portraits can convey physical characteristics, psychological traits, behavioral patterns and information about the physical environment of the subject/user. Therefore, contrary to what happens with portraits that fulfil the function of data doubles, data portraits that act as data mirrors are based on the deliberate collection of personal data, made through service providers and personal devices. Often using statistical analysis of data, data mirrors present more diversified visualization strategies, oscillating between objectivity and subjective expression. This diversity is also reflected in the output formats, sometimes explored as complementary expressions of the same data portrait. For example, a static image can evoke aspects of traditional portraiture, such as the crystallization of a moment in time, materialized for future contemplation. In contrast, dynamic visualizations are evolving self-portraits that unfold in real-time, seeking to highlight the changing nature of human experience and identity over time.

Fig. 5. Characteristics of the data portrait as a data mirror.

CONCEPTS			MECHANICS		EXPERIENCE	
Function	*Theme*	*Content*	*Data collection*	*Analysis*	*Surface*	*Dynamics*
DATA MIRROR	**Type of portrait** Self-portrait **Motivation** Internal	**Data source** Virtual environment Real world **Typology** Behavioural data Biometric data Environment data **Ethnographic value** Physical characteristics Psychological traits Behavioural patterns Physical environment	**Data collection** Service providers Devices **Level of control** Deliberate **Recording format** Textual Numerical	**Data analysis** Statistical None	**Informational structure** Relational Temporal Spatial Spatio-temporal **Representational stance** Analytical Expressive **Modes of expression** Visual Audiovisual Material **Output format** Static image Physical object Moving image Digital interface	**Behaviour of the output** Contemplative Interactive **Nature of the output** Static Transient Variable

Finally, when devised as *critical discourse* on the quantification of all aspects of human experience as digital data, data portraits often point towards the various political implications of the current uses of personal data (Fig. 6). Consequently, they source their data from virtual environments taking advantage of the behavioural surplus, and commenting on its commodification, by extracting as much information on the individual as possible. Hence the broad ethnographic scope and diversity of data collection processes, also tackling possible ways of reappropriating data traces from service providers, devices of everyday use, and web 2.0 infrastructures. Similar to the reflexive functions of data portraits that act as data mirrors, these critical artifacts also explore diverse visualization strategies to reappropriate personal data and use it to the user's advantage, as a tool for self-knowledge. Consequently, this kind of critical stance can be internally or externally motivated, having both applied or artistic purposes, and combining a critical discourse with the role of a self- or individual portrait.

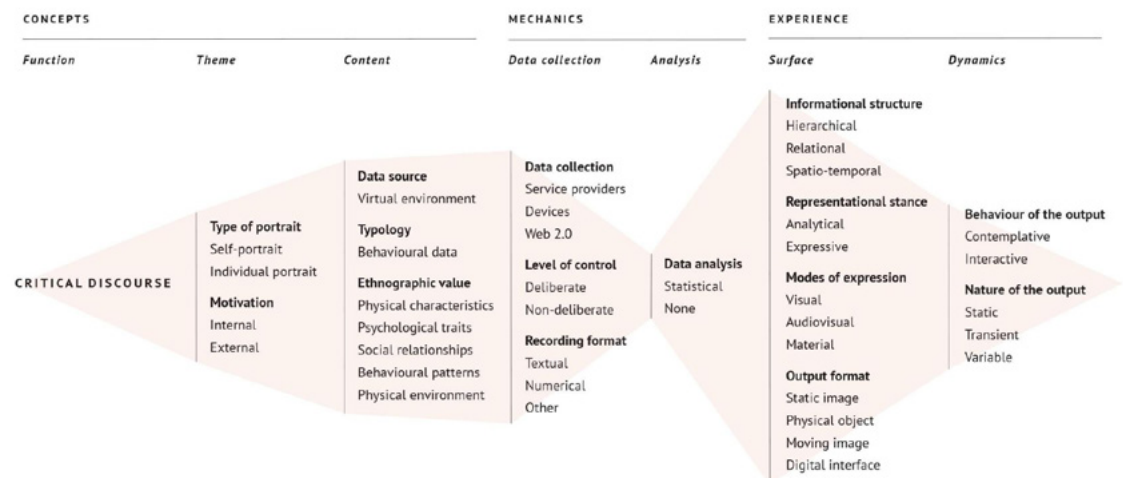Fig. 6. Characteristics of the data portrait as critical discourse.

CONCEPTS			MECHANICS		EXPERIENCE	
Function	*Theme*	*Content*	*Data collection*	*Analysis*	*Surface*	*Dynamics*
CRITICAL DISCOURSE	**Type of portrait** Self-portrait Individual portrait **Motivation** Internal External	**Data source** Virtual environment **Typology** Behavioural data **Ethnographic value** Physical characteristics Psychological traits Social relationships Behavioural patterns Physical environment	**Data collection** Service providers Devices Web 2.0 **Level of control** Deliberate Non-deliberate **Recording format** Textual Numerical Other	**Data analysis** Statistical None	**Informational structure** Hierarchical Relational Spatio-temporal **Representational stance** Analytical Expressive **Modes of expression** Visual Audiovisual Material **Output format** Static image Physical object Moving image Digital interface	**Behaviour of the output** Contemplative Interactive **Nature of the output** Static Transient Variable

One of the aspects differentiating these functions are the sources and typology of data, as well as their ethnographic value as expressed by the visualizations (Fig. 7). Even if these functions of data portraits are better accomplished when more data sources and typologies of data are articulated, as ethnographically relevant data for conveying identity, relevance is also relative to the context in which the data portrait is produced and circulates.

7. Conclusion

This analysis sought to point out the design strategies and methods that become relevant to the communicative and expressive functions of data portraits, as designed artifacts that take advantage of the potential of using personal data as a raw material to represent and convey identity, as well as to critique and raise awareness of the implications of our digital footprint, and how it reveals identity traits.

To fulfil these communicative and expressive functions, data portraits can use the same sources of data but diverge in their visualization strategies. When the data portrait is designed according to a *reflexive* or a *proxy* function, it usually resorts to relational and spatio-temporal informational structures, since it favours the perception of data variability over time. In addition, it also makes clear the human provenance of the input data, as in *Halo* (2017), by Peter Crnokrak. However, when the data portrait aims at establishing a critical discourse on the appropriation of personal data by service providers or questions the lack of user agency over their data, the visualizations tend to be devised as metaphors of the big data universe which, according to Ge, et al. (2018), is characterized by its volume of data, variety, and veracity. In this case, the visualizations usually favour a contemplative experience, through a cumulative representation of the different types of data that are generated and collected about individuals, as with *Poisonous Antidote* (2016), by Mark Farid.

In sum, these visualizations of personal data are designed to convey relevant identity traits, both for the represented individual and within their social

sphere. At the same time, they also question the dematerialization and technological mediation of social interactions, by dwelling on the dissociation between our physical and our digital selves.

Consequently, data portraits also point to a reconceptualization of portraiture, informed by the creative potential of computational media, and reflecting changes in the way we create and experience this representational genre. In this sense, we can evoke how, thirty years ago, in the conclusion of his essay on *Portraiture* (1990), Brilliant wrote that maybe portraiture as a distinctive genre would disappear when "instead of an artist's profile portrait the future will preserve only complete actuarial files, stored in some omniscient computer, ready to spew forth a different kind of personal profile, beginning with one's Social Security number" (Brilliant 1990, 174). On one hand, this statement evokes the extension of current practices of data extraction for profiling individuals, which we can observe in the context of surveillance capitalism. On another hand, it also suggests how the creative practices addressed in this paper can take advantage of the potential of personal data to convey personal identity in ways that more directly correspond to one's lived experience than what mere appearance can convey.

Therefore, data portraits can be seen as a counterpart to Brilliant's idea, by taking advantage of personal data as a raw material for portraiture, distanced from an analytical stance and favouring subjective expression. As a means of visualizing subjectivity, and as visualizations of a subjective nature, data portraits point towards an expansion of portraiture, rather than its disappearance. They suggest how portraiture evolves as "numbers are infiltrating the last redoubts of the personal" (Wolf 2010) and this representation genre expands becoming more attuned to our current mode of living immersed in data.

References

Almeida, Pedro.
2017. *Visactivism: A Visualização de Informação na Perspectiva do Design Activista*. Lisboa: Repositório da Universidade de Lisboa.

Bolter, Jay David, and Richard Grusin.
2000. *Remediation: Understanding New Media.* Massachusetts: MIT Press.

Bonde, Allen.
2013. "Defining Small Data", *Small Data Group.* <https://smalldatagroup.com/2013/10/18/defining-small-data/>.

Brilliant, Richard.
1990. *Portraiture*. London: Reaktion Books.

Ciuccarelli, Paolo, Giorgia Lupi, and Luca Simeone.
2014. *Visualizing the Data City: Social Media as a Source of Knowledge for Urban Planning and Management.* Milan: Springer.

Donath, Judith.
2001. "Mediated Faces", *Sociable Media Group.* <http://smg.media.mit.edu/papers/Donath/MediatedFaces/MediatedFaces.CT2001.pdf>.

Donath, Judith, Alex Dragulescu, Aaron Zinman, Fernanda Viégas, Rebecca Xiong, and Yannick Assogba.
2014. "Data Portraits", *Sociable Media Group.* <http://smg.media.mit.edu/papers/Donath/DataPortraits.Siggraph.final.graphics.pdf>.

Donath, Judith.
2017. The Social Machine - *Designs for Living Online.* Massachusetts: MIT Press.

Floridi, Luciano.
2011. The Informational Nature of Human Identity. In *Minds and Machines*, Volume 21, Issue 4. <https://doi.org/10.1007/s11023-011-9259-6>

Foucault, Michel.
1988. "Technologies of the Self", In Martin, Luther H., Huck Gutman, Patrick H. Hutton (eds.) *Technologies of the self: A seminar with Michel Foucault.* Amherst: The University of Massachusetts Press. (pp. 16-49).

Fry, Ben.
2008. *Visualizing Data.* Sebastopol, CA: O'Reilly Media, Inc.

Ge, Mouzh, Bangui, Hind Bangui, and Barbora Buhnova.
(2018). "Big Data for Internet of Things: A Survey". *Future Generation Computer Systems*, Volume 87, October 2018. (pp. 601-614).

Goodyear, Anne C., Jonathan Walz, Kathleen Campagnolo, and Dorinda Evans.
2016. *This Is a Portrait If I Say So: Identity in American Art, 1912 to Today.* Yale: Yale University Press.

Harary, Yuval Noah.
2015. *Homo Deus: A Brief History of Tomorrow.* New York: Harper Collins Publishers.

Haraway, Donna.
1991. *Simians, Cyborgs and Women: The Reinvention of Nature.* London: Free Association.

Hernández-Ramirez, Rodrigo.
2017. Technologies of the Self: How Are Digital Tools Affecting Human Ontologies?, In *Proceedings of xCoAx 2017: 5th Conference on Computation, Communication, Aesthetics & X*, (pp. 68–80). <http://2017.xcoax.org/xCoAx2017.pdf>

Joler, Vladan, Andrej Petrovski, Djordje Krivokapic LL.M, Dušan Ostraćanin, Nikola Kotur, Jan Krasni, Kristian Lukic, Miloš Stanojević, Olivia Solis Villaverde, and Tamara Pavlovic.
2015a. "Invisible Infrastructures: Mobile Permissions", *Share Lab* <https://labs.rs/en/invisible-infrastructures-mobile-permissions/>.

Joler, Vladan, Andrej Petrovski, Djordje Krivokapic LL.M, Dušan Ostraćanin, Nikola Kotur, Jan Krasni, Kristian Lukic, Miloš Stanojević, Olivia Solis Villaverde, and Tamara Pavlovic.
2015b. "Invisible Infrastructures: Data Flow", *Share Lab* <https://labs.rs/en/invisible-infrastructures-data-flow/>.

Kennedy, Helen, Thomas Poell, and José van Dijck.
2015. Data and Agency. In Big Data and Society. In *Sage Journals*, Volume: 2, Issue: 2. <https://doi.org/10.1177/2053951715621569>

Lee, Catarina.
2014. *Uma Exploração Criativa da Noção de Transmutabilidade: Do Conceito à Prática.* Lisboa: Repositório da Universidade de Lisboa.

Levin, Golan.
2009. "Audiovisual Software Art: A Partial History", *Flong* <http://www.flong.com/texts/essays/see_this_sound_old/>.

Lupi, Giorgia.
2017. Data Humanism, The Revolution will be Visualized. <http://giorgialupi.com/data-humanism-my-manifesto-for-a-new-data-wold>

Lupton, Deborah.
2014. "Self-tracking cultures: Towards a sociology of personal informatics", In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design. (*pp.77–86). <https://doi.org/10.1145/2686612.2686623>.

Lupton, Deborah.
2016. "You are Your Data: Self-Tracking Practices and Concepts of Data", *Lifelogging - Digital Self-tracking and Lifelogging – Disruptive Technology and Cultural Transformation.* Berlin: Springer VS. (pp. 61-79).

Lupton, Deborah.
2016. *Quantified Self.*
Cambridge: Polity Press.

Manovich, Lev.
2002. "Data Visualization as New Abstraction and Anti-Sublime", Lev Manovich's Personal Website. <http://manovich.net/index.php/projects/data-visualisation-as-new-abstraction-and-anti-sublime>.

Manovich, Lev.
2013. The Computer as Metamedia. In *Software Takes Command.* New York: Bloomsbury Publishing. (pp. 103-113).

Medeiros, Margarida.
2000. *Fotografia e Narcisismo: O Auto-retrato Contemporâneo.* Lisboa: Assírio e Alvim.

Meirelles, Isabel.
2013. *Design for Information.* Massachusetts: Rockport Publishers.

Morley, David.
2007. *Media Modernity and Technology – The Geography of the New.* New York: Routledge.

Mun, Min, Shuai Hao, Nilesh Mishra, Katie Shilton, Jeff Burke, Deborah Estrin, Mark Hansen, and Ramesh Govindan.
2010. Personal Data Vaults: A Locus of Control for Personal Data Streams. In *Proceedings of 2010 ACM Conference on Emerging Networking Experiments and Technology,* CoNEXT 2010. DOI: 10.1145/1921168.1921191.

Pink, Sarah, Robert Willim, and Melisa Duque.
2018. "Broken Data: Conceptualizing Data in an Emerging World", *SAGE Journals.* Doi/full/10.1177/053951717753228.

Poletti, Anna, and Julie Rak.
2014. Identity Technologies: Constructing the Self Online. Wisconsin: University of Wisconsin – Madison.

Renaud, Alain.
2003. "O Interface Informacional ou o Sensível no Seio do Inteligível". (Cruz, Maria Teresa. trans.). Louise Poissant. (org.), *Esthétiques des Arts Médiatiques, Interfaces et Sensorialités.* (pp. 65-89). Québec: Presses Universitaires du Québec. (Original em Francês, n.d.).

Rettberg, Jill Walker.
2013. *Seeing Ourselves Through Technology: How We Use Selfies, Blogs and Wearable Devices to See and Shape Ourselves.* New York: Palgrave Macmillan.

Ribas, Luísa.
2014. "Digital Computational Systems as Aesthetic Artifacts: Practices, Systems, Processes and Perspectives", *Proceedings of xCoAx International Conference on Computation, Communication, Aesthetics & X 2014.* [accessed: 12-03-2017] <http://2014.xcoax.org/pdf/xcoax2014-Ribas.pdf>.

Sampaio, Catarina, Luísa Ribas, and Pedro Ângelo.
2019. "Data (self) Portraits: An Approach to the Visualization of Personal Data from an Autoethnographic Perspective". *Proceedings of xCoAx 2019: 7th Conference on Computation, Communication, Aesthetics & X.*

Sampaio, Catarina, and Luísa Ribas.
2020. "Representing the Self Through the Visualization of Personal Data". *International Journal of Creative Interfaces and Computer Graphics (IJCICG)12(1).* DOI: 10.4018/IJCICG.2021010101.

Sampaio, Catarina, and Luísa Ribas.
2020. "On Data Portraits: An analysis of creative approaches to the visualization of personal data". *Proceedings of EVA London: 30th International Electronic Visualization and the Arts Conference.* DOI: 10.14236/ewic/EVA2020.46.

Selke, Stefan.
2016. *Lifelogging - Digital Self-tracking and Lifelogging – Disruptive Technology and Cultural Transformation.* Berlin: Springer VS.

Stepanchuk, Yulia.
2017. Quantified Construction of the Self: Numbers Narratives and the Modern Individual. *In Proceedings of the International Conference Internet and Modern Society.* (pp. 28–36). https://doi.org/10.1145/3143699.3143720

West, Shearer.
2004. *Portraiture.* Oxford: Oxford University Press.

Whitehead, Tony L.
2005. "Basic Classical Ethnographic Research Methods", *CEHC Cultural Ecology of Health and Change.* <http://www.cusag.umd.edu/documents/workingpapers/classicalethnomethods.pdf>.

Whitelaw, Mitchell.
2008. "Art Against Information: Case Studies in Data Practice", *The Fibreculture Journal.* <http://eleven.fibreculturejournal.org/fcj-067-art-againstinformation-case-studies-in-data-practice/>.

Wolf, Gary.
2010. *"The Data-Driven Life", In The New York Times, 2 April 2010.* <https://www.nytimes.com/2010/05/02/magazine/02self-measurement-t.html>.

Xiong, Rebecca, and Judith Donath.
1999. "PeopleGarden: Creating Data Portraits for Users", *Association for Computing Machinery.* <https://dl.acm.org/citation.cfm?id=322581&dl=ACM&coll=DL>.

Zuboff, Shoshana.
2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* London: Profile Books Ltd.