National and Kapodistrian University of Athens

Interdepartmental Graduate Program

"Science, Technology, Society—Science and Technology Studies"

Department of History and Philosophy of Science

&

Department of Informatics and Telecommunications

MSc Thesis

Title: Picturing Big Data in media: an STS-based approach

Name of Student: Anastasia Stoli

Registration Number: 13/218

Thesis Advisory Committee:

Tympas Aristotle, Professor (advisor)

Tzafestas Elpida, Associate Professor (member)

Varoutas Dimitris, Associate Professor (member)

June 2019

Abstract

Big Data is one of the key emerging technologies. Scholars have started to study the ways textual content concerning Big Data affects the public's response to the technology but have yet to study the effect images might have on the matter. This study aims to start filling this gap in the study of Big Data literature through research on Big Data imagery in newspapers in Greece and the U.K. Using a combination of content analysis and pictorial semiotics, I have placed Big Data images in newspapers under five main themes: blueness, dataveillance, big data-related anxiety, gender construction and analogue/digital. The thesis relies heavily on the interdisciplinary field of study known as Science, Technology, Society or Science and Technology Studies (STS).

Key words: big data, science communication, science and technology studies, content analysis, pictorial semiotics, dataveillance, gender bias, analogue vs digital, liberal vs conservative, printed media, digital doubles

Table of Contents

Introduction

Out of all the emerging technologies belonging to the Fourth Industrial Revolution, Big Data may be the most pervasive one. It is everywhere and, at the same time, nowhere. Part of all facets of everyday life but somewhat elusive, undetectable. It is a technology seamlessly fused into other entities, lurking in the background, operating tirelessly in the shadows. Present behind social media platforms, multinational corporations like Google and Amazon, amassing facts and figures via closed-circuit television (CCTV) and even calculating the odds of success in famous sports, such as baseball, through the use of analytics. With the expansion of digital economy, Big Data additionally serves as a panacea for organizations who wish to digitize their business, or position themselves in the explosive market structured around modern, internet-mediated human activities. Utilizing servers with high processing capabilities, adept at analyzing countless bytes of data and deciphering them into valuable information, appropriate Big Data action is taken on behalf of public or private organizations in order to generate profit or sustain/advance politics (Degli Esposti, 2014, p.210).

Concurrently, Big Data functions in a way that unavoidably leads to several potential concerns. The astounding amount of control Big Data allows leads to its exploitation by a plethora of actors. Either politically, as Snowden's release of NSA documents in 2013 showcased, or economically, such as the more recent 2018 Facebook–Cambridge Analytica data scandal did (Cadwalladr & Graham-Harrison, 2018).

And yet, it seems that despite its rather disconcerting nature Big Data is not only accepted by the public but valorized by the academic community (Smith, 2018; Van

Dijck, 2014) for its insight and intuition, much higher than any human could possibly offer. Ultimately, Big Data is presented as both the perpetrator, aiding in acts of crime, and the saviour, the knight in shining armour, offering the solutions needed to encounter the issues it has formerly created (Smith, 2018, p.8).

There exists a rich bibliography concerning Big Data and its relation to the public, either positively or negatively (Mayer-Schönberger, 2014; Heldbing, 2015). Moreover, there is scholarly work available on the rhetoric surrounding Big Data, either in popular or scholarly texts, belonging to newspapers and academic journals accordingly (Papacharissi, 2015; Dourish & Gómez Cruz, 2018). All references of Big Data, be it by scholars or journalists, inform and educate the public regarding the technology. Thus far, however, most of the work in the field has been concerned with studying the ways this technology is communicated to the public through words. To my knowledge, with the exception of a very recent paper by Pentzold, Brantner and Fölsche, published in 2019 under the title "Imagining Big Data: Illustrations Of "Big Data" In US News Articles, 2010–2016", no other work has been written about the ways Big Data is *visualized* to the public. This gave me an extra reason to choose to undertake research on the ways in which Big Data is visualized in images published as accompaniments to popular articles written in popular UK and Greek newspapers. The only other work similar to mine that I have managed to find, deals with the visualization of Big Data in US News articles, and has consequently addressed the gap in research there. I chose to not focus solely on my own country, Greece, not only because a sample originating from two countries, Greece and the UK, would be considerably richer, but also because focusing on Greece-only runs the risk of producing idiosyncratic results with no additional impact on the academic

community outside our borders, should it turn out to be a "singular case". I have chosen UK newspapers as the second source, due to their influence within Europe. Published in English, they are accessible internationally.

Visuals are powerful, images tend to exercise a considerable influence in shaping attitudes towards everything (Allen, 2019). Such influence can be used to effectively communicate ideological messages. In the case of Big Data, that might mean taking something discrete or abstract and turning it into something personal, alarmingly impactful for the society and sensational enough to generate revenue for the news agencies. Indeed, as my research will show, imagery of Big Data points to both bad and good elements of the technology at hand, providing visuals that evoke utopian and dystopian scenarios (Pentzold, Brantner, Fölsche, 2018, pp.2-4). Visuals are also key in detecting gender bias in technology communication (Tympas, Konsta, Lekkas and Karas (2010). Unlike the articles, which oftentimes go through extensive overview to avoid including stereotypes and misconceptions, images are more likely to be overlooked. And since most of the images used by journalists in articles belong to stock photography, which traditionally contains many stereotypes (Frosh, 2001), gender prejudices are more prevalent and adhere to a continuity, an established link between Third Industrial Revolution advertisements and Fourth Industrial Revolution advertisements.

I began preparing this thesis with a list of research hypotheses. They were the following three:

1. Images visualizing Big Data evoke predominantly negative emotions to the public, regardless of the content of the article they accompany.

2. Images visualizing Big Data contain unconscious gender bias which exhibits similarities to advertisements of computers belonging to the Third Industrial Revolution.
3. Liberal newspapers use negative imagery to communicate Big Data more than their conservative counterparts.

By the end of my empirical research, my first two hypotheses were confirmed, whereas the third one was partly confirmed while simultaneously providing an interesting secondary layer, which will be discussed later. I also discovered more findings that I had not anticipated in my original three hypotheses. To carry out my empirical research, I based my approach in the STS sub-field of Science Communication, using content analysis, paired with pictorial semiotics, as my methodological tools.

I collected my empirical data by entering the internet archives of the newspapers I chose, retrieving images accompanying articles relevant to Big Data and examining them both qualitatively and quantitatively, with a focus on the former rather than the latter. I will be presenting a more detailed description of my methodology in a separate chapter.

Chapter 1: Theoretical Framework

1. What is STS?

In this short chapter, I will provide a brief account of STS as a field of studies, relying heavily on Sergio Sismondo's "*Science and Technology Studies and an Engaged Program*", published in MIT Press in 2007. According to Sismondo, STS complements but also challenges traditional perspectives in philosophy, sociology, and history of science and technology. It critically addresses policy, governance and funding issues of relevance to science and technology. STS' objective is the reform of science and technology in the name of equality, welfare and the good of the environment (Sismondo, 2007, p.13). Sismondo (2007) believes that "constructivist STS has created a space for theoretically-sophisticated analyses of science and technology in explicitly political contexts" (p.13). Social constructivism posits the idea that science and technology are highly social, decidedly active and very much "not themselves natural". Should I attempt a standard history of STS, I could begin with Thomas Kuhn's "Structure of Scientific Revolutions" (1962) which emphasized a communal basis of the solidity of scientific knowledge. After Kuhn's work followed Bloor's (1976) and Barnes' (1974) "strong program" in the Sociology of Knowledge. Participants in the strong program wished to start from a naturalist explanation of scientific knowledge and from that point and on, investigate causes of knowledge. The strong program provided STS with a theoretical backdrop for studying the construction of scientific knowledge. Feminist STS, another important field within the disciplinary, showed how ideology can also contribute to the construction of scientific knowledge. Lastly, the Empirical Program of Relativism (EPOR), mostly due to

Harry Collins' work in the 1970s, bore much similarity to the strong program and its set of values and goals (Sismondo, 2007, p.14).

Learning from the study of controversies is an important aspect of STS. In the 1970s, a number of researchers, most predominantly Harry Collins, Karin Knorr Cetina, Bruno Latour, Michael Lynch and Sharon Traweek, simultaneously adopted a novel approach of studying cultures of science. They moved into laboratories in order to see "science in action". For them, phenomena are constructed in laboratories and what is found in them is not nature but a result of human effort (Sismondo, 2007, p.15). According to Sismondo (2007), "laboratory phenomena are not in themselves natural but are made to "stand in" for nature; in their purity and artificiality, they are typically seen as more fundamental and revealing of nature than the natural world itself is" (p.15).

Trevor Pinch and Wiebe Bijker's work in 1987 transferred concepts from the study of science to the study of technology, under the title "Social Construction of Technology". Abbreviated to SCOT, their position argued that the success of a technology depends on the strength and size of the groups that take it up and promote it. Actor-Network Theory (ANT), a similarly chief theory from within STS, was introduced by Bruno Latour and Michael Callon within the same decade. One could pose that ANT combines the interpretive frameworks of EPOR and SCOT with the materialism of laboratory studies (Sismondo, 2007, p.16). In the words of Sismondo (2007), "the metaphor of construction, in its generic form, thus ties together much of STS: Thomas Kuhn's historiography of Science, the strong program's rejection of non-naturalist explanations, SCOT's interpretive flexibility of straightforward technologies and ANT's distribution of agency of techno-science" and even though "these programs are not unified, yet, the metaphor has enough

substance to help distinguish STS from more general history of science and technology" (p.17).

As a field of studies, STS is concerned with both the available science and technology but, also, with the promotion of a socially responsible science. In recent STS work the constructivist program has extended itself to public sites and dealt with cases that combined the theoretical framework with public interests. Perhaps the most important aspect of STS is that it does not separate scientific and political practices, choosing instead to study them as inseparable, providing a genuine and holistic study of societies where knowledge and technology are not "externalities to political processes" (Sismondo, 2007, pp. 18-19, 21, 26). This is why I find STS to be most fitting for studying Big Data visualizations and how they affect the public. The field is theoretically equipped to handle both the epistemic part of what Big Data is and the social consequences of their very existence. It is important to note however, that STS is a field of studies brimming with theories. Some established from scholars within, others originally belonging to other fields of knowledge but adopted and oftentimes expanded by STS researchers. For this reason, in the following chapters I will be presenting key STS theories, as well as theories, beyond STS, which are concerned with the visualization of science and technology. I will be explaining their traits and methods and how they could potentially be utilized for the analysis of my empirical data. Ultimately, I will choose only one of them for this research. I will begin with two theories which are considered the most renowned in STS literature and history, Actor-Network Theory and Social Construction of Technology. Then, I will move on to smaller-scale theories, such as Frame Analysis or theories from the sub-field

Science Communication, which grew out of a sub-field called Public Understanding of Science.

1.2. Leading STS Theories

1.2.1 Social Construction of Technology

In this chapter, I will be providing an outline for one of the leading theories within the STS field, Social Construction of Technology. To do so, I will be utilizing the work of Trevor J. Pinch and Wiebe E. Bijker (1984, 2009) as well as that of Langdon Winner (1993), who formulated some criticism on the subject. SCOT is a constructivist approach to studying science and technology that emerged in the 1980s. It posits the idea that the truth of scientific facts and the working of technical artifacts are constructed socially. Additionally, it adopts the single artifact as its unit of analysis (Bijker, 2009, p.88). It is an amalgamation of two schools of thought: the empirical program of relativism and the social constructivist approach to the study of technology. For SCOT, technology does not develop autonomously. Whether or not a technology is accepted has no basis on epistemological grounds or the natural world. (Pinch&Bijker, 1984, pp.401) Instead, "it is all social" (p.402). The central concept of SCOT is the so-called 'technological frame', which parallels Kuhn's paradigm to a certain degree, with the exception that the 'technological frame' can be applied to all kinds of relevant social groups whilst Kuhn's paradigm was exclusively intended for scientific communities. Furthermore, regarding the role of expertise in public debates, SCOT suggests that all "relevant social groups" have some form of knowledge. Ultimately, SCOT amounts to a concrete rejection of all teleological, linear and one-dimensional approaches to technological development (Bijker, 2009, pp.89-93).

Given that the artifact is SCOT's preferred unit of analysis and society, a three-step heuristic approach is offered. In the first step, the artifact is described through the eyes of relevant social groups. Identification of these groups is achieved by pinpointing which actors refer to (perceive) the artifact in the same manner. This is where a second key aspect of the first step comes into play, the researcher's demonstration of the "interpretive flexibility" all artifacts inevitably possess (Bijker, 2009, p.90). As interpretive flexibility lessens, some artifacts gain dominance over others and the various meanings of an artifact converge into a singular one. During this process of social construction, there is a sense of "closure" and "stabilization" as one artifact establishes itself whilst the others weaken and wither away in its shadow (Bijker, 2009, p.90). The third and last step is an attempt to analyze and explain why and how a specific artifact dominates the rest and is henceforth crowned "a success story". This is where SCOT highlights a legitimate difficulty in all attempts to write down the history of technology. Most analyses concerning key technological events in human history focus on successful, accepted technological innovations and not rejected ones, ultimately providing an asymmetrical account of what has happened (Pinch&Bijker, 1984, p.405).

Moreover, SCOT offers a conceptual framework meant to uncover the hidden political dimensions of technological culture. As Bijker states in his 2009 paper on "*Social Construction of Technology*", "[t]echnology is socially and politically constructed, society including politics is technically built; technological culture consists of sociotechnical ensembles." (p.92). The sheer multitude of technologies in modern society poses serious difficulties for anyone who seeks an overarching grasp of human experience in a technological society. A social constructivist mode of inquiry recommends that rather than

employing broad-gauged assumptions that reproduce technological determinism, scholars need to talk more precisely about the dynamics of technological change. SCOT looks very closely at the artifacts and the social actors whose activities affect their development (Winner, 1993, pp.362, 364).

Social constructivists wish to open the "black box" of historical and contemporary technology to see what is inside. The term "black box", refers to a device or system that is described solely in terms of its inputs and outputs, what is going on inside need not be understood (Winner, 1993, pp.365). SCOT reveals a rich spectrum of technological choices, alternatives, and branching points within patterns sometimes thought to be necessary (Winner, 1993, pp.366). According to Winner (1993), "contingency and choice is emphasized over forces of necessity in the history of technology" (p.367).

SCOT's approach has been criticized by scholars like Langdon Winner as narrow, due to the ignorance of important questions about technology and human experience. There is an almost total disregard for the social consequences of technical choice. For Winner, consequences are seldom a focus of study, with SCOT being unconcerned with what the introduction of these new artifacts means for people's sense of self, the human communities adopting them, the changes in power within society and the effect on the quality of everyday living. The focus in SCOT's study of technology lies on the origins and dynamics of technological innovation (Winner, 1993, p.368).

Choosing which social groups are relevant is also problematic in the sense that oftentimes there are groups that have no voice but will be affected by the results of technological change nonetheless. These groups could have been suppressed or deliberately excluded from the overall process, relegating their possibilities in the realm of

nondecisions. Social scientists who do not take into account this type of power play, where groups are consistently excluded from power, essentially offer a political and societal account that is implicitly conservative. SCOT escapes the bind of Whig history, in which the past is described as a sequence of steps unavoidably leading somewhere better, to the accomplishments of today, but still falls victim to elitism and seems to have nothing to say about the deep-seated political biases that underlie the spectrum of choices that surface for relevant social actors (Winner, 1993, pp.369-370).

Furthermore, SCOT scholars don't seem to evaluate the morality or political nature that technologies possess. The premise of interpretive flexibility works well in cases in which social consensus is achievable but in circumstances when there are serious disagreements regarding the technological innovation, it does not apply as satisfactorily. SCOT's methodological grounds do not allow for a theoretical or practical position on technology and human well-being, which means that no values are associated with any technology being studied. Social constructivism provides no solid, systematic standpoint or core of moral concerns from which to criticize or oppose any particular patterns of technical development (Winner, 1993, pp.371-374).

While SCOT surveys the evidence, it offers no judgment on what it all means, besides noticing which technological projects succeed and which fail, how new forms of power arise while others decline. (Winner, 1993, p.375) In the words of Winner (1993), "choices are available, the course of technological development is not foreordained by outside forces and is instead a product of complex social interactions" (p.375).

1.2.2. Actor-Network Theory

In this chapter I will be providing a short overview of Actor-Network Theory, referencing Bruno Latour's book "Reassembling the social", as well as notable works by Collins, Yearly (1992) and Cressman (2009). Actor-Network Theory (ANT), whilst intimidating and oftentimes difficult to comprehend, is nonetheless an exceptionally widespread model, utilized across several disciplines. Its origins are mainly associated with Michael Callon, Bruno Latour and John Law. Methodologically, in Latour's own words, Actor-Network Theory is concerned with "science and technology in the making" and is indifferent to the "ready-made science and technology". ANT explicitly rejects technological determinism and all possible linear models of scientific and technical change. It performs micro-level studies of the places where science and technology come into being, traversing from laboratories to institutes and funding agencies (Cressman, 2009, pp.1-3).

Network builders, typically engineers and scientists, serve as the primary actors in an ANT approach. Their standpoint is used to interpret the process of network construction. The tracing of the complex relationships between actors, be it governments, technologies, knowledge, texts, money or people, effectively opens the black box of science and technology. But what makes ANT differ from SCOT or other sociotechnical approaches? When defining an actor, ANT considers texts, humans and machines alike. A non-human element can be an actor within a network and should be afforded the same analytical and descriptive framework as a human element. Technologies, a non-human element, can encompass both political, social and economic elements, as well as science, engineering and the particular histories of these practices. As generalized symmetry would posit, since

from an empirical point of view both technologies and humans play equally important roles in construction of actor-networks, there is no point in establishing a duality between them. Though heterogeneous, their identities as actors are defined by their interaction with each other and the associations they form describe how some networks come to be larger and more influential than others (Cressman, 2009, pp.3-4, 10).

As Latour himself says in his book "*Reassembling the social*" published in 2005, "[t]he social is bound together with participants who are members of a 'society', and are called 'social actors', non-social entities can become participants as well. A systematic tracking of these participants and the relationships between them ends up in a shared definition of a common world, a 'collective' (p.247). The process of technical innovation itself effectively cancels out any pre-existing divide between technical, scientific and social context. The actors within the framework continuously redefine the sociotechnical world, making ANT differ further from SCOT, an approach mainly concerned with studying stable social groups in order to explain the meanings technical objects have. For ANT, what matters is a symmetrical account of the social and non-social, a reciprocal relationship between them, in an effort to describe the reasons and ways the technologies we "invent", have come to be. As Cressman puts it: "We delegate to technologies the work of many humans and in turn, the technologies delegate behavior back onto the social." Until the mid-1990s, ANT was utilized within STS but from that point onward it moved beyond STS and is now considered to be a widely applicable theory of the social world (Cressman, 2009, pp.5, 8-10).

ANT treats dichotomies with consistency. Natural "actants" and human actors are to be treated symmetrically (Collins & Yearly, 1992, pp.309-10). By extending symmetry,

humans lose their pivotal role and nonhuman actants' power needs to be calculated properly. ANT generalizes symmetry by treating all actants that are party to the scientific enterprise in the same manner (Collins & Yearly, 1992, pp.321-3). According to Collins and Yearly (1992), "while [ANT] is philosophically radical as a model, its usage is essentially conservative, a poverty of method making it subservient to a prosaic view of science and technology (p.323).

1.3.2. Frame Analysis

In this chapter, I will be providing a short description on Frame Analysis, based on secondary literature by Pan, Kosicki (1993), as well as Benford, Snow (2000) and Maier, Rothmund, Retzbach, Otto & Besley (2014). Since the 1990s, framing processes have come to be regarded, alongside resource mobilization and political opportunity processes, as a central dynamic in understanding the character and course of social movements (Benford, Snow, 2000, p.612). Erving Goffman, in his 1974 book “Frame analysis: an essay on the organization of experience”, as quoted in Pan, Kosicki (1993), supports that we all “actively classify, organize, and interpret our life experiences to make sense of them”. These "frames," as he characterizes them, enable individuals "to locate, perceive, identify, and label" occurrences or information. Todd Gitlin’s 1980 book, entitled “The whole world is watching: Mass media in the making and unmaking of the new left”, links the concept of framing directly to the production of news discourse. As referenced by Pan and Kosicki (1993), Gitlin argues that this procedure "enable[s] journalists to process large amounts of information quickly and routinely [and to] package the information for efficient relay to their audiences". William A. Gamson, similarly quoted by Pan and Kosicki (1993) states that a frame is a "central organizing idea or storyline that provides meaning" to events related to an issue. This can either be a metaphor, an exemplar, a catchphrase, a depiction, or, as is often the case, a visual image (Pan, Kosicki, 1993, p.56).

According to Benford and Snow (2000), framing analysis is an “active, processual phenomenon that implies agency and contention at the level of reality construction”. It is active because something is being done, and it is processual because its process is ever so dynamic and involving (p.614). This kind of agency is attributed to organizations, activists

and oftentimes traditional media journalists, who decide which scientific issues are placed onto the public "agenda" and how they're "framed." In doing so, as Maier, Rothmund, Retzbach, Otto & Besley (2014), they "effectively shape the information the public may learn and what attitudes it may develop towards them as a result" (p.88). Framing, therefore, deals with the question of how journalists depict scientific information. They seem to "strategically choose an 'organizing idea', a so-called frame, and their decision lends greater weight to certain considerations and arguments over others" (p.90). How they choose to frame a certain technological innovation for example, translates it into either a problem or an opportunity, it assigns blame or praise to whoever is responsible for it and as a final point, offers suggestions on how the public should respond to it (Maier, Rothmund, Retzbach, Otto & Besley, 2014, p.90).

Framing Analysis offers a very appropriate framework for a research like mine. It is a framework that is frequently used in connection to the study of issues like the one I am considering in this thesis. One of its key domains of use has to do with looking into how traditional journalism "frames" certain themes in order to evoke specific emotions or attitudes to the public. In the case of Big Data, this could imply an intentional, conscious decision to present them as something negative, daunting or unfamiliar so that the public responds to the sensationalism this type of journalism arouses. What would furthermore explain this decision would be the fact that Big Data is already part of the public's everyday life. Just by using a phone or a computer, or even by crossing a street fitted with camera equipment, Big Data enters people's lives. The personal computer had to "create" a market and convince the public to welcome it into their everyday life and because of that, the framing of computers when they were first introduced was deliberately warm and inviting

(Sumner, 2012). By contrast, journalists have the freedom to frame Big Data as they wish, even negatively, with no fear of collapsing a new market.

1.3.3. Science Communication

In this chapter, I will introduce a sub-field of STS, called Science Communication, which grew out of a sub-field called Public Understanding of Science (PUS). I will be basing my presentation on previous works by Wynne (1995), Weigold (2001), Bultitude (2011) and Short (2013). PUS, PAwS (Public Awareness of Science) or even SiS (Science in Society) are all descriptive titles given to a late 20th century movement by governments, industry, scientists and various stakeholders. The Royal Society's 1985 report *The Public Understanding of Science*, otherwise known as the Bodmer report, offered the first description of this sub-field, one that concerned itself with the proper education of the public regarding science (Short, 2013, pp. 39-40). From then on and for a substantial amount of time, PUS was equated with public appreciation and support of science, and with the public's "correct" understanding and use of "technical" knowledge and advice. In this approach, a deficit model was established, which assumed that the public lacked necessary knowledge about scientific concepts/theories. These scientific concepts/theories had to be properly communicated from the scientists to the public (Bultitude, 2011, p.5). The PUS research agenda, according to Wynne (1995) was thus confined to "measuring, explaining, and finding remedies for apparent shortfalls of correct understanding and use, as if it was free of framing commitments that have social implications" (p.4).

PUS offers three distinct research approaches for identifying and measuring public attitudes towards science: large-scale quantitative surveys meant to measure the levels of public scientific literacy, reconstruction of "mental models" the public has regarding scientific processes and scientific content and lastly, qualitative field research, interested in the public contextualization of scientific expertise, exploring how people in different

social contexts experience and construct its meaning (Wynne, 1995, p.6). The deficiency model suggests that widespread ignorance regarding science can be reduced if the public is educated about it. The model also supports that the public's attitude will become positive as a secondary result (Weigold, 2001, p.173).

The deficiency model would eventually come to be discredited by a number of studies that showed that supplying more information to people does not necessarily change their views. Evidence has systematically pointed towards a resistance in accepting new ideas from an assortment of ethical, religious and cultural beliefs. And yet, it has also been noted that people who opposition scientific knowledge actually possess greater levels of understanding of science than the non-attentive, average public which means that being educated about science does not necessarily foster a positive scientific attitude. Some kinds of knowledge produce anti-science sentiments and sometimes, generalized positivity around science diminishes once more specific scientific issues come into play (Weigold, 2001, p.186). The notion that public opinion can be easily swayed by objective evidence is a "fallacy" (Short, 2013, p.41). The deficiency model's most important problem has been its top-down, science-centered approach. What historically followed the deficiency model was the "dialogue model", which required an understanding of the context of scientific knowledge and how different people utilized it in their individual lives (Weigold, 2001, p.174). This new model framed the public as someone who should be listened to, someone who possessed a meaningful input into scientific policy making. Public engagement refers to specialists listening and interacting with non-specialists (Short, 2013, p.41).

Unlike PUS, Science Communication, which in its broad definition encompasses public engagement's "two-way communication", recognizes the mutual learning that

occurs by both publics and scientists during their interactions. It is a process during which all participants learn, not just the public (Bultitiude, 2011, p.1). Within Europe, the majority of citizens feel that scientists cannot be trusted to tell the truth about scientific issues due to their dependence on industrial and private funding. Furthermore, they distrust the validity of science itself because of the ever increasing press coverage of controversies (such as climate change, nuclear power and genetically-modified food) and public disagreements between respectable scientists. The emergence of alternative experts operating outside the scientific community and the scandalous publications of scientific fraud, successful or otherwise, have additionally reduced the public's trust in science and scientific authority (Bultitiude, 2011, p.1).

Science Communication is an important enterprise for any nation, for it provides the public with technical skills and knowledge that not only improve their personal lives but also improve the state's overall economic outcome as it creates a significantly more capable workforce. At the same time, Science is both a major part of a country's culture, as well as an influential "player" in its central politics. Science education is a way to share this heritage with its citizens which will then be able to utilize this knowledge to participate in major democratic decisions. For Science Communication, a dialogue approach is adopted, which involves a two-way information exchange between the public and the scientific community. The scientists might still possess the scientific facts but the public has an equally important, alternate understanding of issues in need of solving, which stems from their personal interests, particular experiences and local knowledge (Bultitiude, 2011, pp.4-5).

Given the theme of this thesis, Science Communication seems to be the most relevant sub-field to draw approaches from. The visualization of Big Data images accompanying articles in major UK and Greek newspapers actively educate and shape public opinion on this emerging technology. At the same time, the public's reaction to the contents of these newspapers, affect in turn the technology and the community surrounding it. Personal experiences regarding the technology are communicated from the public to the scientific community, public and private organizations, as well as the government, usually through the use of media, such as TV, newspapers and online blogs.

1.3. Big Data

In this chapter, I will introduce definitions of Big Data and issues connected to the social use of Big Data. I will do so by relying mostly on articles published in a set of STS and related ournals, which include *Science, Technology and Society*, *Public Understanding of Science*, *Big Data and Society*, *Osiris* and *New Media & Society*. Most of the articles were retrieved electronically from the official webpages of each respective journal, whilst some were found in digital libraries of academic journals such as JSTOR and SAGE Journals.

First, I will provide a brief historization of the term *data* to prevent some common pitfalls, like presentism and technological determinism when discussing the Big Data "revolution". Afterwards, I will continue with defining what Big Data is and how it works. Thirdly, I will present human and non-human actors working within the Big Data framework and how they help enrich the databases, either voluntary or involuntarily. Finally, I will reflect on some of the public's positive and negative outlooks regarding the technology itself and why the public seems to simultaneously embrace and fear Big Data's capabilities and characteristics.

A common account of the emergence of Big Data is that the field in question is a product of late 20th century technological leaps, formed by the improved infrastructure that emerged out of the massive use of electronic computers. This account presents Big Data as a transformative force for modern scientific practice and epistemology. But most technologies have histories that challenge this kind of linear narration of progress. In the case of Big Data, its history arguably goes back to a period well before electronic

computers were available. Societies have dealt with an overabundance of information numerous times in the past, be it printed works in medieval times, observations of animals during the European expansion or even bureaucratic accumulation of numbers in the 19th century. There has been, however, an undeniably dramatic technological innovation after the advent of electronic computers, which has accelerated and amplified the features of data-driven science. Computers have irrefutably changed the way data is stored, processed and interpreted. Unlike the pre-electronic era, today's Big Data radically transcends the circumstances and locality of its production (Aronova, von Oertzen, Sepkoski, 2017, pp.2-7, 15-16).

Even if, historically speaking, Big Data is but a chapter in a long history of observation and quantification, for the purpose of this thesis, I will be concerned with Big Data as a field defined solely within the context of the Third and mainly the (so called) Fourth Industrial revolution. As such, the term describes information that has been counted, quantified, digitized and reduced to binary electronic signals stored in colossal digital datasets (Aronova, von Oertzen, Sepkoski, 2017, pp.1, 6) which are routinely exploited to create new forms of value, using innovative analytical techniques. (Pentzold, Brantner, Fölsche, 2018, p.2) It refers to the sheer scope of modern information technology that is available to us, from Google analytics to NSA surveillance, from biodiversity databases and genomics archives to the interactions of subatomic particles (Aronova, von Oertzen, Sepkoski, 2017, p.1).

At the same time, the term "Big" is used not only to describe the sheer size of these datasets, but also the vast scale of analysis required to "tackle them on" (Pentzold, Brantner, Fölsche, 2018, p.2). A project this big and daunting needs major investments of

time, energy, personnel and capital and as a result, whenever a decision regarding funding and media attention is made, the size and effort behind Big Data possess remarkable "weight". Data-driven science has effectively created a new cultural and political landscape. Within Big Data, there are two types of personal data; data sets (or doubles) and data aggregates. Data doubles are bound to individual users whilst data aggregates are generated from several individual inscriptions (Aronova, von Oertzen, Sepkoski, 2017, pp.8-9). 'Datafication' is the name of the process by which information gathered regarding the public is translated into "discrete, machine-readable, measurable, manipulable bits and bytes" (Hildebrandt, 2013, p.6).

Big Data is generally produced via interactions with online technologies. These interactions can be between human and nonhuman actors, but also between two or more nonhuman actors. The contribution *netizens* provide to these immense datasets is accomplished by a vast array of activities, such as call-making, usage of phone applications, online information searches and uploads of personal content (textual or visual) on social media platforms, such as Facebook (Michael, Lapton, 2015, p. 1). The reasoning and motivation behind all the data-sharing is the formation of social experiences and even the construction of the user's identity. This kind of -often-times- excessive data-sharing is further encouraged and normalized by the digital infrastructures and platforms themselves. Social media is built in a way that instills in its users an impulse to self-reveal and to voyeuristically spectate the lives of others. The manner in which people view and contribute content online makes them both agents and subjects of surveillance. The socio-material infrastructure of the Web and the predominantly visual culture permeating modern

society significantly affect netizens so that they desire to perform in ways that enrich the field of Big Data (Smith, 2018, pp. 9, 11-12).

Contribution can also be made in complete absence of human actors, as is the case of the *Internet of Things*. Continuous generation of digital data is achieved via self-tracking devices human actors have equipped on themselves but are not actively or consciously using, CCTV cameras surveilling the public on a 24-hour basis and *smart homes*, whose domestic appliances produce data concerning the house occupants and share them with the "cloud". Looking at Big Data as a sociotechnical assemblage, its responsive, dynamic nature is apparent in the way it constantly reconfigures itself using new data generated by both human and non-human actors, combining them in relentlessly different ways (Michael, Lapton, 2015, pp. 1-2, 4).

Research done regarding the public's opinion on Big Data has brought forth feelings of deep ambivalence and occasional suspicion. Although the value of Big Data is recognized for it maintains national security, helps control crime, promotes public health and improves healthcare, at the same time questionable surveillance activities, leaked documents, various hacker attacks on databases containing private and personal information (such as medical data or the information they have uploaded to adult dating sites) which are then manipulated by social media companies (e.g. Facebook) and sold for commercial benefit to third parties, have demonstrated to the public just how insecure the technology is. By sharing credit card details online to purchase a product, or posting a photograph on Facebook, people consent to their data being collected. But in the case of machines sharing to other machines people's online behaviors, be it public transportation behaviors or biometrical behaviors, we have a fall out of the data protection legislation.

And what might be the biggest concern due to the obscure nature of the activity, are inferred data, profiles derived from data mining that may have the biggest impact on an individual. Their profile might stop them from getting the job they want, the education they wish for or even raise the price of their insurance (Hildebrandt, 2013, p.10). People organize their lives around data and due to their ease of use, their convenience and stimulating properties, they forget their multidimensional properties and their capacity to track them in varying ways (Smith, 2018, p.10).

Beneficiaries of data-driven environments are security and profit orientated organizations, be it state agencies or businesses. In today's knowledge economies, personal information assemblages are valuable commodities (Smith, 2018, p.3). Because of this, the more the public realizes the commercial value of their data doubles, the more they exhibit hostility towards the idea that government agencies sell data for profit instead of using them for the public good (Michael, Lapton, 2015, p. 5, 7). Cloud computing technology, for example, offers flexibility, easy set-up, efficiency and cost benefits. And yet, while affordable, accessible and convenient in the services it provides, cloud consumers remain aware of potential threats, such as hackers, spammers, terrorists, data breaches, and businesses using people's data for financial gain. These data privacy issues are further exacerbated by an ambiguous legislation regarding cloud computing (Abed, Shavan, 2018, p.4, 12). Techno-pessimism says that today's society is on the verge of a data dictatorship and its people are becoming incapable of perceiving reality outside the mediation of Big Data-driven techniques and technologies (Hildebrandt, 2013, p.7).

The current explosion of data actually turns them, to a large extend, into noise; the sheer quantity of bits and bytes makes them unreadable to the human eye (Hildebrandt,

2013, p.6). Repeated encounters with digital data operate to naturalize these entities, blindsiding the public to the ways they get implicated in exploitation and governance. Power and capital increasingly transfer through data and yet people generally possess only a limited awareness of how data exercise influence over their lives in important, if often in obscure and unseen ways. 'Data doxa' is about the way digital data have come to be perceived in Western societies as normal, necessary and enabling. 'Dataism' is becoming a widespread ideological belief where the enlightening, emancipatory and optimizing properties of digital technologies are accentuated and where greater supply and accumulation of information is thought to reveal truths about the natural and social world. An explanation of the public's simultaneous misunderstanding and dependence on data boils down to what Smith (2018) calls 'data doxa', a "sensibility that limits critical engagement with data beyond the immediate ends they serve" (pp.1-5).

Chapter 2: Methodology

Science Communication utilizes three main forms of media: traditional journalism, face-to-face events and online interactions. For the purposes of this thesis, I will be focusing on traditional journalism, which possesses a number of advantages and disadvantages as a medium. Printed journalism in the form of newspapers has the advantage of reaching a wide audience, especially if the newspaper in question enjoys high popularity or has been historically established as a legitimate source of information. In this context, it is also important to note that newspapers have always been and still are considered an agenda-setting medium, whose handling is done by professional journalists, who impart on it high quality and general trustworthiness. Conversely, scientists rarely control the way journalists will cover their work which is usually given a superficial or limited focus due to the very nature of printed media and their inherent prerequisite to sell, leading to a prioritization of 'sensationalist' stories. The format of a newspaper article also limits the interaction to a one-way street, from the journalists to the public (Bultitude, 2011, pp. 7-8).

For this portion of the Methodology, I will be presenting my primary sources, the purpose I chose them for, the reasoning behind these specific choices and the way the data collection needed for this thesis was achieved. For primary sources, I decided to use the online editions of two British newspapers and three Greek ones. I chose to focus my research on the digital copies of these newspapers because, even though traditional media such as newspapers are still an important source of information for the public regarding science and technology, the Internet is quickly "leaping ahead" as a tool for procuring knowledge and answering questions (Maier, Rothmund, Retzbach, Otto & Besley, 2014,

pp.86). While newspaper circulation has unavoidably declined in the last few years, demographically popular newspapers continue to reach new audiences online (Pentzold, Brantner, Fölsche, 2018, p.5-6). In fact, UK's Office for National Statistics (ONS) has calculated a dramatic increase in Britons using the Internet to inform themselves. From 20% of adults reading newspapers and magazines online in 2007, to 55% in 2013 (Office for National Statistics, 2013), 62% in 2015 (Office for National Statistics, 2015) and 64% by 2017 (Office for National Statistics, 2017). Lastly, utilizing the digital versions of these newspapers effectively solves one of the limitations of printed media as a medium for Science Communication: in electronically-published articles, the public can not only read the content of the text but also comment underneath, turning a previously one-way communication to a two-way communication.

The newspapers I will be using for my research purposes will be: *Kathimerini (Καθημερινή), Proto Thema (Πρώτο Θέμα), EFSYN/Eleftherotypia (Εφημερίδα των συντακτών/Ελευθεροτυπία), The Guardian and The Telegraph*. *EFSYN* and *Eleftherotypia* refer to the same newspaper, whose title was changed in 2014 after a short period of intermittent publication. All five newspapers enjoy high circulation numbers in their respective countries. *Kathimerini* and *The Telegraph* are categorized as conservative right-wing, *Proto Thema* as populist right-wing, *EFSYN/Eleftherotypia* and *The Guardian* as liberal and left-wing. I find it important to note that websites oftentimes produce duplicates, results of articles that have been de-published or pictures that appear unavailable. Due to the issues stated above regarding the calibration of these websites, some articles or photographs from the period of years I will be covering might still be missing from my

final sample. That aside, here is how the research for the visualization of Big Data was done.

Beginning with *The Telegraph*, I collected all images published in the online edition of the newspaper, spanning from December 2012 to March 2019. I searched for Big Data, using the website's search engine, the keywords "+Big +Data", typed in this manner to help the site's algorithm with the results and lastly, the option "relevancy" to discard as many unrelated results as possible. The newspapers' online archives resulted in a total of 1700 articles. I went through them, spending a few minutes on each article to ascertain whether the phrase Big Data existed inside and the item was not placed in the results due to a website malfunction. I was interested in determining whether Big Data was an important part of each article, accompanied by a picture that visualized the technology and not just a throwaway line that should not be counted in this research. When the results started becoming irrelevant, I continued reading the abstract for the rest of the result pages but stopped entering and reading the entire article. What remained after this process was a sample of 35 images visualizing Big Data, which I will use for my analysis. The largest amount of images belong to articles published in 2016 and 2017.

Proceeding with *The Guardian*, I counted on a similar process. I collected all images published in the online edition of the newspaper, spanning from March 2012 to March 2019. I searched for Big Data, using the website's search engine, only unlike the case of The Telegraph, The Guardian's website algorithm did not respond positively to the "+Keyword A +Keyword B" format, so I typed a simpler "Big Data" keyword search instead. The newspapers' online archives resulted in a total of 316 articles. I went through all of them, spending a few minutes on each article to ascertain whether the phrase Big

Data existed inside and was an important part of content, accompanied by a picture that visualized Big Data and not some other theme from within the article in question. After this was complete and I had excluded all the articles that did not address Big Data-related issues explicitly or did not contain relevant images, what remained was a sample of 148 images visualizing Big Data that I will use for my analysis. The largest amount of images belong to articles published in 2013 and 2014, perhaps as a result of the Edward Snowden leaks affecting the public sphere and media at the time (Pentzold, Brantner, Fölsche, 2018, p.7).

Moving on to the Greek media, I will start with *Kathimerini*. I collected all images published in the online edition of the newspaper, spanning from February 2014 to January 2019. Similar to the British case, I began by locating the website's search engine and typing the keywords "+Big +Data" to assist the site's algorithm in finding the most relevant articles and choosing the "relevancy" option to prioritize them appropriately. My search resulted in 107 articles. I read all articles, spending again a few minutes on each, to ascertain whether the phrase Big Data existed inside and the item was not placed in the results due to a website malfunction. Furthermore, I was interested in determining whether Big Data was an important part of each article, accompanied by a picture that visualized the technology and not just a throwaway line that should not be counted in this research. Upon completion of this procedure, I collected a sample of 17 images visualizing Big Data, which I will use for my analysis. The largest amount of images belong to articles published in 2014.

The second Greek newspaper I researched was *Proto Thema*. I collected all images published in the online edition of the newspaper, spanning from May 2012 to January 2016.

I located the website's search engine and typed the keywords "+Big +Data" to assist the site's algorithm in finding the most relevant articles. Additionally, I chose the subcategory 'Technology' that the website provided for more specific results and the option 'relevancy' to improve the search. My search resulted in 212 articles, which I read in total, spending a few minutes on each to ascertain their relevancy and find the appropriate pictures, wherever they existed. After I excluded all the articles that did not address big data-related issues explicitly or did not contain relevant images, what remained was a corrected sample of 10 images visualizing Big Data that I will use for my analysis. Despite the significantly vast original pool of results, the Big Data-related images published in *Proto Thema* are a rather small sample.

Last of all, I researched a third Greek newspaper whose name was changed throughout the years of interest to this research. It was named *Eleftherotypia* until 2014 and *EFSYN* from 2014 until now. The results from both websites will be analyzed conjointly. I collected all images published in the online edition of both newspapers, spanning from May 2013 to March 2019. I located the websites' search engines and typed the keywords "+Big +Data" to assist the sites' algorithms in finding the most relevant articles. The format "+KeywordA +KeywordB" proved useful in the case of Eleftherotypia where it successfully filtered out irrelevant results but was deemed unnecessary for the *EFSYN* website where typing "Big Data" was enough for the engine to produce only appropriate results. The Eleftherotypia website search resulted in eight articles while the *EFSYN* website search resulted in 25. After reading all 33 articles and downloading the pictures visualizing Big Data, what remained was a sample of 19 images that I will use for

my analysis. Spanning from 2013 to 2019, there was not a specific year to note where a notably bigger amount of articles regarding Big Data was published.

The subsequent archive of visual representations of Big Data contained 46 pictures from the Greek newspapers and 183 from the British newspapers, totaling in 229 pictures as the complete sample. I organized them both by content and date, tagging the various themes each of them communicated to the viewer. For the purposes of this thesis, I chose a combination of content analysis and pictorial semiotics. Content analysis is either qualitative or quantitative and usually refers to the breakdown of text data, which is coded into explicit categories and is either described using statistics or contextual meanings and themes (Hsieh, Shannon, 2005, p.1278). In my case, due to the pictorial nature of my sample, I replaced text data with images, which serve the same purpose. My sample is photographs and pictorial advertisements used to compliment articles on Big Data and pictorial semiotics is the method of analysis utilized to understand the meaning of the pictures involved (Sonesson, 1993, p.3).

I have chosen a summative approach to content analysis, which typically refers to a qualitative analysis that begins with quantifying data (in this case, pictures), based on their content in order to understand their usage and infer their meaning. The quantitative aspect of the procedure is limited to the counting of the existing themes within the sample. And although I will briefly mention this quantitative aspect, my purpose is to move beyond it and embark in latent content analysis, in which I will interpret these recurring patterns to unearth underlying meanings (Hsieh, Shannon, 2005, pp.1283-4). Offering one keyword as an example, the first step is calculating the amount of times pictures showing

dataveillance appear in my sample and then, interpreting what they mean, based on existing secondary literature.

The strength of such approach is that it is systematic. In content analysis, the entire sample is taken into account, no cases are chosen just to "prove a point" and exceptions to the motifs at hand are similarly analyzed so their legitimacy can be integrated into the overall analysis. Additionally, content analysis is flexible enough that it can be used by a variety of disciplines and researchers. It can also be combined with other methods of analysis, as I have done in this thesis. The added value of the pictorial semiotics included in this approach is the sheer richness of the field. Images brim with ideological constructions and rich connotations, whose veiled meanings are conveyed to the viewer, oftentimes in a subliminal manner (Cox, 2009, p.92).

The weaknesses of this approach lie in the interpretation of the image by the analyst. The examination is concerned solely with the content of the image and not the process by which it came to be. There is an additional, subjective, human factor in the production, choice and publication of images in newspapers which is entirely excluded from the research process. Moreover, whilst flexibility aids in the coalition of methods, it also makes content analysis vague enough that scholars sometimes find it difficult to incorporate or use in their research. The lack of a firm characterization regarding what it entails and how it works essentially limits its application (Hsieh, Shannon, 2005, p.1278). Ultimately, I have chosen this combined method despite its disadvantages and because it fits well with my line of enquiry (Cox, 2009, p.92).

Chapter 3: Findings-Interpretations

To begin the analysis of my empirical data, I first completed a quantitative research to calculate key motifs within the sample. I will be presenting here, both in text and in an information table the themes with the highest percentages found in the 229 images visualizing Big Data existing within my sample. In this chapter, I will provide some numerical figures, beginning with the highest percentages and moving on to the lowest ones. In the subsequent chapters, I will be analyzing each of these motifs in depth, offering critique based on both existing secondary literature by distinguished scholars, as well as my own personal assessments. I will be accompanying the evaluation with select images from the sample.

To begin with, a staggering 35% of the overall sample (77 out of the 229 images) are images whose colour is homogeneously and overwhelmingly blue. Following this, there are two themes that fall under similar categorization. Together, amounting to 31% of the total sample, they can be described as Big Data-related anxiety. Separately, 17.5% of this Big Data-related anxiety (40 out of the 229 images) is visualized in the form of either geolocation tracking fears or shadows of people signifying the emergence of data subjects. The other 13.5% (31 out of the 229 images) is visualized in the form of privacy concerns related to dataveillance.

Moving on, a significant percentage of the sample, not only due to its size but also due to the diversity of meanings contained within it, is the representation of gender. 16.6% of the sample (38 out of the 229 images) displays biased representation of gender in ways eerily akin to preceding advertisements of technological objects, such as computers.

Finally, 9% of the sample (21 out of the 229 images) attempts to visualize Big Data themselves. This visualization codifies Big Data as either analogue or digital. Taking a step back from content analysis, and after having examined the sample, I will be attempting a

secondary reading of the results, combining qualitative and quantitative analysis to show a possible correlation between the political affiliation of the newspapers in question and the positive or negative visualization of Big Data, as well as an overall comparison of Greece and the U.K.

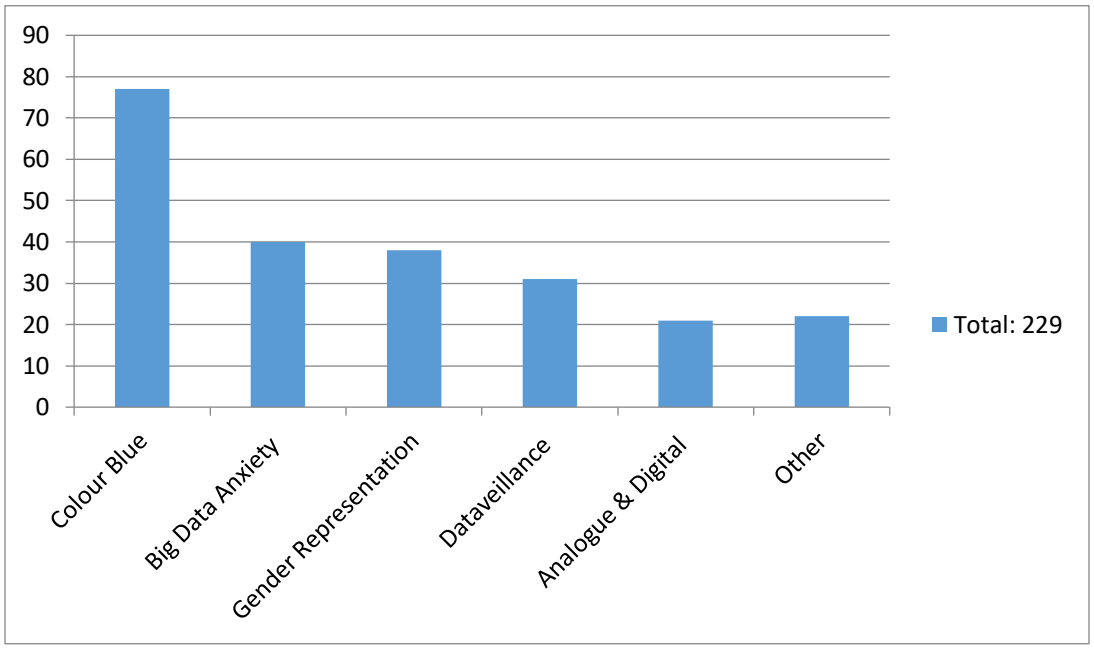


Table 2 showcases the number of images in which each major theme within the images concerning Big Data appears in all five newspapers of my sample, from largest to smallest.

3.1. Blueness

A noteworthy 35% present of the sample (approximately 77 out of the 229 pictures) is comprised of images which are heavily tinted with the colour blue. Utilizing colour theory and behavioral psychology, I will endeavor to place some significance in its use when portraying Big Data. There are two chief, generally-accepted, traditional colour theories, the RYB Model (Red-Yellow-Blue) and the RGB Model (Red-Green-Blue). These models suggest that three primary colours exist, either red, green and blue, or red, yellow and blue. In either case, the interaction with each other subsequently produces all other colours, which are called secondary and tertiary (Hirsch, 2004).

The characteristics these three primary colours possess are their hue, intensity and saturation. The modification and combination of these three characteristics eventually lead to the creation of secondary colours, such as orange, purple and more (Aslam, M., 2006, p.17). An additional division is oftentimes made between warm and cool colours which seems to have its origin in description of art back in the 18th century and is possibly inspired by weather patterns. In which case, rainy or cloudy weather, which tends to be overwhelmingly blue in hue is described as ‘cool’, whilst a sunny, clear sky, which tends to ‘bask’ in colours, which are products of the primary red, such as orange, is described as ‘warm’ (Goldsmith, 1764). And since summer, whose weather is predominantly warm has longer days, whilst winter whose weather is predominantly cold, has shorter days, the Purkinje shift phenomenon leads to the human eye perceiving more blue tones in the dark and more red tones in the light, ultimately leading to an association being made between coldness and blue and warmness and red (Colman, 2008). Already, there seems to be psychological associations and meanings given to specific colours or groups of colours.

In relating colour to human behavior, a distinction must be made between meanings people attribute to colours instinctively and meanings they attribute to colours as a result of learned associations. Because of the different meanings and perceptions which are given to colours from one culture to another, it is more likely that colour associations are a result of past experiences, religious beliefs, art, literature and language (Aslam, M., 2006, p.18) rather than a physiological or biological response. In psychology literature, studies are made in which colors and types of emotions are systematically linked and these

associations remain consistent across studies, meaning that empirically-speaking, there is concrete evidence of a systematic relationship between color, emotions and psychological functioning (Labrecque, Milne, 2011, p.713) But what does this mean for the visualization of Big Data predominantly in blue colours?

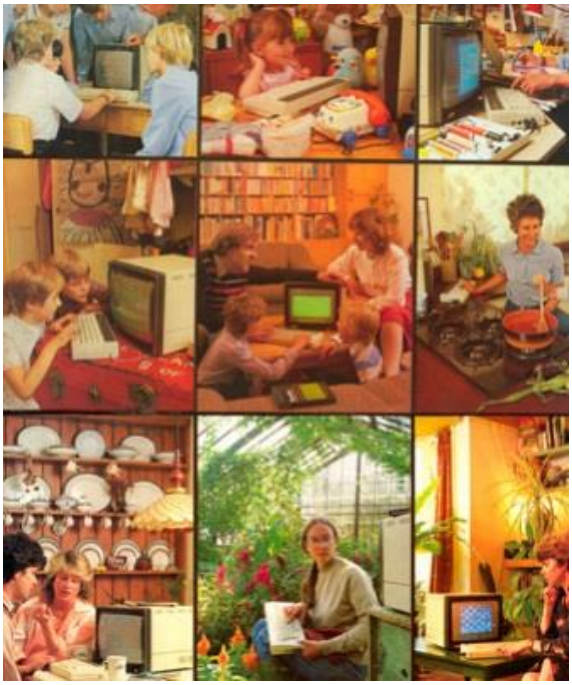
[1] First vertical row are images from the Guardian, second vertical row are images from The Telegraph, third vertical row are images from Proto Thema, fourth vertical row are images from Kathimerini and fifth vertical row are images from Eleftherotypia/EFSYN.

From the 20th century until today, and situated strictly within the geographical borders of the West, blue, with very few exceptions, represents masculinity. Anglo-Saxon countries in particular, such as the U.K. and the U.S., which are of interest to this research, not only consider blue masculine, but also high quality and corporate. Though it is important to note that for the U.K., masculinity is also denoted by red, blue is used

overwhelmingly to signify stability, responsibility and the "corporate man" (Aslam, M., 2006, pp.19, 25).

Because of concerns regarding both the vague nature of Big Data and the security and privacy concerns that have arisen, especially after the Snowden report, it is important that Big Data is visualized in a way that inspires safety and trust. Blue has been consistently associated with intelligence, communication, efficiency, duty, and logic and is often seen as the colour of choice for famous brands who routinely deal with data gathering, such as Facebook or Twitter. The decision is not accidental, for the apparent competence of a brand depends on the presence of this specific hue (Labrecque, Milne, 2011, p.714). The trust and security the blueness of these logos or images inspire, is not a warm, friendly trait that one could relate to domestication or a maternal sense of safety and closeness. Instead, it is a kind of trust closer linked to security, a special kind of trust that one has for structures and institutions. It inspires reliance and confidence but it remains cold to and detached from the person experiencing it.

A suggestive paper by Sumner (2012), regarding the domestication of computers as an emerging and somewhat confusing technology that entered the market in the 1970s and 1980s, shows vintage advertisements that were published in magazines and newspapers in hopes of creating a need for computers where none was before, such as in image [2]. These ads recreated moments of domesticity, utilized the idea of the nuclear family and placed computers inside the comforting confines of *home*. The objective was to connect computers to everyday life so that the public would feel motivated to purchase them, even if they were unaware or confused as to their practicality (Sumner, 2012, p.1). At the same time, it was vital that the image of the computer to the general public would shift from intimidating, "cold" and feared machines used by military to "warm", empathetic family staples. The image of the family essentially transformed computers into socially-friendly machines and pushed away fears of privacy and other concerns the public



[2] Old Computer advertisements from Sumner's personal collection as published in "Today, Computers should interest everyone", 2012, p.7.

might have had. The technology was turned from a "frightening icon into a domesticated friend of the family" (Reed, 2000, pp.174-178).

This apparent effort to familiarize the public with home computers was not only achieved via the use of familiar settings, nods to improved education for the children and parallels to already-established and useful home appliances, but also via the use of a specific colour wheel. In fact, if one looks at the advertisements Sumner has published alongside his article [image 2], warm colours are readily apparent, such as red, orange and yellow. I posit that warm colours were used for their psychological associations. Red is associated with excitement and stimulation, emotions fitting for introducing a new technology into your everyday life and it is also considered up-to-date, similarly to the product it is used to promote. Likewise, orange is considered sociable and lively, generating feelings of familiarity and homeliness (Labrecque, Milne, 2011, p.714). With these in mind, brands such as IBM most likely intentionally chose these colours to positively affect their brand and up their prospective sales. But why wasn't the same thing done to familiarize Big Data as an emerging technology to a somewhat dubious and uninformed public?

Big Data, as shown in image [1], is visualized in rich blue hues. Even in the percentage where blue was not the primary colour in the image, warm colours were still noticeably absent. Coupled with the recurrent appearance of shadows of people and a sense of alienation, both of which I will be analyzing in a later chapters, blue provides an additional psychological connotation. Not only is it a colour that inspires trust, stability and masculinity, but also a colour that further aids the viewer in experiencing estrangement from the technology, because of the specific type of impersonal trust it denotes. Looking

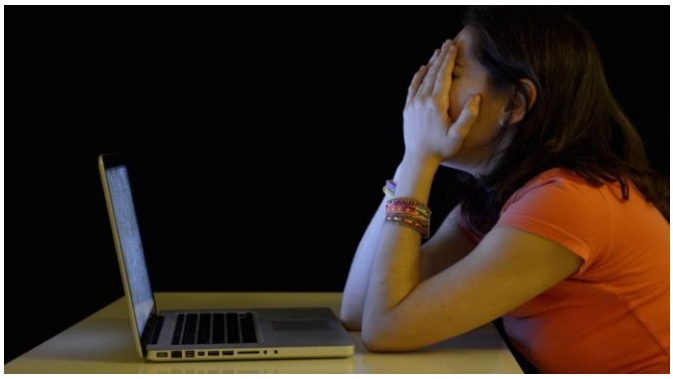
at images visualizing Big Data, there is a genuine impression that their nature is uncertain to the general public and hard to pinpoint. Even though the public currently uses Big Data in their everyday life significantly more than the public in the 1970s and 1980s used computers, the visuals tell a different story. Whereas the warm images of computers in the 1970s and 1980s brought the public closer to them, the cool images of Big Data furthers its distance from them. But given the concerns regarding dataveillance and privacy breaches that are ever-present, there seems to be a trade-off; an either conscious or unconscious decision to decrease the sense of familiarity in order to increase the sense of security.

3.2. Dataveillance

Another recurrent theme in the visualization of Big Data is dataveillance. A noteworthy 13.5% present of the sample (approximately 31 out of the 210 pictures) is comprised of images which symbolize dataveillance, either in a more traditional imagery or with the use of digital cameras [image 4] and drones [image 5]. As Degli Esposti (2014) notes, mass dataveillance is the "systematic monitoring of people or groups, by means of digital information management systems, in order to regulate or govern their behavior" (p.211). Such activity oftentimes incites feelings of domination due to the fundamentally asymmetrical relationship between the ones who control the data-gathering, be it a company or the government and the employee or customer who has no option but to accept being surveyed should they want to participate in any given "transaction" (Degli Esposti, 2014, pp.211, 215).

Typical methods of surveillance differ from dataveillance in the regard that the goal of dataveillance is to capture as much data as possible about everything, endlessly, and with no plans of ever letting go of the massive cumulative information housed in colossal servers (Andrejevic, Gates, 2014, p.185). Unlike traditional surveillance, which collects information for particular uses, and, generally has a specific "object" of surveillance, dataveillance unceasingly amasses immense amounts of data for several unstated preset purposes, of interest to a variety of public and private institutions which are, more often than not, reticent about their involvement (Van Dijck, 2014, p.205). A typical, high-tech surveillance method is the drone, capable of capturing an impressive amount of data, including all available wireless data traffic in any area it is monitoring. The drone presents a technology capable of performing both traditional surveillance, targeted spying, as well as dataveillance, in the form of opportunistic data capture (Andrejevic, Gates, 2014, p.185).

Dataveillance does not solely describe the action of observing and recording the public via cameras, sensors and the use of satellites and drones. It also consists of identification markers that further assist in tracking the public's movements and behaviours, such as credit cards, identification, personalized numeric codes and other electronic passes which monitor daily activities. Then, this aforementioned collected information is transformed into knowledge through the use of analytics. "Raw" data is grouped in search for patterns that have value, either commercial or investigative (Degli Esposti, 2014, pp.211-212). Finally, dataveillance not only passively collects and analyzes everyday activities but uses it to manipulate the public's behaviour [image 3]. Those whose behaviour is targeted and manipulated via dataveillance techniques do not necessarily realize it, and for that reason the term 'manipulation' is oftentimes used. Degli Esposti (2014) describes behavioral manipulation as "the ability of influencing people's actions intentionally, nudging people's behavior in a desired direction" (Degli Esposti, 2014, pp.211-212, 220).



[3] Image from the article: "In social media, we cry in November and get stressed in April", published in November 2016, "Proto Thema"

'Dataism' and 'Datafication' normalize actions otherwise reminiscent of "Big Brother" where the public's personal information is routinely accessed. Companies and government agencies systematically hail Big Data for its usefulness in either protecting the public or building a superior and considerably more suitable market for their needs, but also academics and researchers consider Big Data a panacea for all questions concerning behavioral knowledge (Van Dijck, 2014, pp.198-199). Scholars view 'datafication' as a "revolutionary research opportunity to investigate human conduct". And yet, the neutrality of 'datafication' is based on the notion that those who are responsible for gathering and distributing these enormous datasets are trustworthy and their methods of quantifying said data is objective. Similar to the separation of church and state, the normalization of this type of social monitoring depends on the very independence and integrity all liable institutions have. However, should it be revealed that all these institutions, governmental,

academic and commercial, are interconnected, sharing data with each other, utilizing the same ecosystem, operating in similar ways and on the same level of infrastructure, then this trust in the neutrality of 'Dataism' and 'Datafication' loses its legitimacy (Van Dijck, 2014, pp.206-7).

[4] From the article "Whiskey Tango Foxtrot by David Shafer review – a hair-raising thriller about data theft", published on June 2015, in The Guardian

Ultimately, the very foundations of 'Dataism' were brought down by a single event which effectively catapulted Big Data firmly into the public's eye in 2013, the year in which most Big Data-related images appear in my sample, especially in the U.K. case. In 2013, Edward Snowden brought to the surface privacy and legal violations committed *in camera*, when he exposed all he had to perform on behalf of the intelligence community (Van Dijck, 2014, p.197). A CIA employee and contractor, Snowden disclosed highly classified information concerning the PRISM program, ran by the National Security Agency (NSA) alongside several governments and telecommunication companies, all responsible for running in collaboration numerous global surveillance programs. Global public attention was effectively drawn to vulnerabilities of the system regarding the public's security, as well as the political and security dimensions of state monitoring. Suddenly, the idea of individuals being hacked, harmed by corporations or profiled by investigative agencies became a tangible reality. Connecting to the Internet ceased being the free, anonymous, private utopia it was advertised to be (Smith, 2018, p.5).

These newfound concerns regarding dataveillance seemingly became the "driving force" behind the emergence of a clearly-visualized anxiety, starting in 2013 and remaining present until today, although in a steadily decreasing popularity. Images

[5] From the article "Big data: how to win the numbers game", published on December 2016, in the Telegraph

[6] From the article "Data Journalism is improving – fast", published on July 2013, in the Guardian

[4], [5] and [6], showcase a tangible fear of being surveyed, either by the use of cameras or drones. Since a significant portion of this surveillance is performed online by private companies, such as Facebook or Twitter, Snowden's revelations had a major effect on brands wishing to retain their clientele's trust and continuous usage of their service. That led most companies to publicly denounce their governmental ties or willing collaborations with agencies such as the N.S.A., in an effort to preserve their legitimacy, saying instead that any cooperation between them and other institutions was "forced" (Van Dijck, 2014, p.204).

All images in my sample that were published before 2013 do not show dataveillance or allusions to breaches of privacy, confirming the correlation between the Snowden incident and the emergence of dataveillance anxiety in the overall population. At the same time, although private companies at the time denounced their willing involvement in the scandal, dataveillance imagery persisted in the following years, even if their frequency declined, signifying that this new facet of Big Data "had come to stay". It is also interesting to note, that a year ago, a significant resurgence of dataveillance, this time strictly in the form of social media platforms data-gathering and selling private information for gain, took place. The Facebook–Cambridge Analytica data scandal in early 2018, brought to light by whistleblower, ex-Cambridge Analytica employee Christopher Wylie, revealed how an inordinate amount of personal data was "harvested" from Facebook profiles by the British political consulting firm Cambridge Analytica without their consent and was subsequently used for political advertising purposes (Cadwalladr & Graham-Harrison, 2018).

This new breach of privacy is clearly visualized in my sample with images that heavily utilize the Facebook logo, as well as images of people being "transformed" into data, to showcase the impersonal usage of the human identity for financial and political gain. The latter will also be present in the imagery discussed in the next chapter. In the images [7] and [8] from my sample, published at the very time the scandal first appeared in the media, this visualization is apparent.

[7] From the article "Here's why tech companies abuse our data: because we let them", published on April 2018, in the Guardian

[8] From the article "Big data for the people: it's time to take it back from our tech overlords", published on March 2018 in the Guardian

Given that the way I have approached these images is by quantifying them as clearly negative representations of Big Data, one could undoubtedly question why such decision was made. Some representation of dataveillance, such as the visualization of drones, could go either way, perceived either positively or negatively depending on the individual reading the article at hand. Additionally, some of the images of Big Data I have codified as negative, compliment articles whose content is decidedly neutral, or even positive, such as that of image [5] which discusses the advantages of using Big Data in sports. But according to William Allen's lecture in the National Documentation Centre of Greece (2019) regarding Big Data and migration, even when an article's textual content is not negative, a negative image accompanying it can still induce significant emotional response to the reader. Furthermore, various literature regarding dataveillance (Hildebrandt, 2013; Michael, Lapton, 2015) has analyzed public responses to the use of the technology and regarded it as negative, systematically affecting

people's sense of safety and personhood. For this reason, and due to basing my analysis in these papers, I am quantifying them as negative media.

3.3. Anxiety, fear and impersonalization

Big Data exists within the messy world of technology, defined by the conflicted interests of institutions, governmental power relations and discourses (Pink, Lanzeni, Horst, 2018, p.2). How is trust defined within this context? Trust can be the ability to act in the immediate future with some, if not absolute, certainty. And hope is imagining this future to be grander and more aspirational than the current present. Humans innately trust activities and practices that are based on routine. They feel safe within them, comfortable in their know-how and accomplished when they complete them. The familiarity of the process lends relief to the individual, reassuring them that even if some impromptu event occurs, they can "fall back" into habits and actions they're accustomed to (Leszczynski, 2015, pp.966-7). Leszczynski refers to this as 'ontological security' and it can go as far back as Hume's *problem of induction*, which describes humanity's inherent need to know what the future holds in order to feel safe (Hume, Norton & Norton, 2000).

Due to the vague and uncertain nature of Big Data, as well as the swiftness with which the emerging technology is constantly redefining itself, be it geographically, ontologically or ethically, the public seems to be experiencing anxiety as a result. This can either be a type of "surveillant anxiety", an aspect which was touched upon in the previous chapter with the analysis of dataveillance, or it can be about their sense of self. On the subject of locational data collection, an anxiety-inducing issue lies in

[9] From the article "Data could be the real draw of the internet of things – but for whom?" published on September 2015, in the Guardian

[10] From the article "There is a leftwing way to challenge big tech for our data. Here it is", published on August 2018, in the Guardian

[11] From the article "Analytics and big data: how artificial intelligence could deliver genuine social impact", published on September 2016, in the Telegraph

the collection and disclosure of an individual's data. This "geosurveillance" can be achieved via a large number of applications and digital practices and it affects even those who do not voluntarily participate in the usage of the WEB or mobile applications. Nonetheless, just owning a mobile phone is enough for this type of GPS tracking to commence and differences in devices and mobile companies is irrelevant, for it is all interconnected anyway. As a result, personal locational disclosures cannot be safely and entirely disabled by individual users, either via the phone settings or inside specific applications, particularly by those who are technologically uninformed. The source of anxiety lies not only in the effort to control the flow of their personal data but also in the realization that "exerting such control is effectively futile" (Leszczynski, 2015, pp.966-7).

In my view, this fear and anxiety is visualized in the form of highways [10,11], aerial images and even the typical image of planet earth [9], with "lines" [11] representing the signal of the GPS device or phone that is shadowed, monitored by private companies and the state alike.

And at the same time, this surveillant anxiety affects the public's sense of self. The immense amount of data that phones, computers and other devices upload daily to Big Data servers, extensively reveal private details about the users' identities (Pink, Lanzeni, Horst, 2018, p.3). In an era of uncertainty, data verify and validate who and what people are but it is not only a habit of over-sharing and revealing personal information that "springs" from this uncertainty. A two-way communication unfolds, in which the public informs digital devices of their identity but then, in turn, relies on what the data has generated to "know, optimize and orientate the self." Complex data processes turn familiar, invoke dependency and seduction. They are treated as a means to an end, a source of entertainment or convenience, and their agency is all at once ignored and demanded (Smith, 2018, p.6-7). In this regard, it is not solely a privacy concern, but also an "anxiety of control", where the constant surveillance and data-gathering reduces individuals' control over the information they disclose. It is not only the information they unwittingly offer just by browsing the WEB and using social media platforms that they cannot control, but also how that information can later be utilized by, unknown to them, corporations and governments. As explained in the earlier Big Data chapter (1.4), the networking of different, theoretically independent actors in this pervasive enterprise of never-ending dataveillance practices, essentially guarantees that the integrity and safety of personal data simply cannot be maintained (Leszczynski, 2015, p.979). The main drive of

[12] From the article "Smart Copenhagen: Can big data improve quality of life?" published on October 2016, in the Telegraph

this secondary aspect of Big Data anxiety, is the fear that the self is constructed online, not by the individual, but by Big Data itself and is in fact, a misrepresentation of the real self (Pink, Lanzeni, Horst, 2018, p.3).

The public's "digital doubles" are assembled, drawn from the users' online behaviour which involuntarily discloses things about themselves. Social networking sites, social media platforms and the WEB in general, code many aspects of social life in a way that has never been quantified before—relationships forged, casual interests, every-day conversations, google searches, locations visited, emotional responses, and so on (Van Dijck, 2014, p.198). The sheer complexity of the human identity turns into an algorithm, an amalgamation of every movement performed online. Predictive analytics then collect and discern all online activity and generate the digital self, down to its political beliefs, sexuality and psychosocial profile (Leszczynski, 2015, p.968).

A substantial amount of my sample shows people bathed in shadows, photographed from the back or "drowning in a sea" of Big Data, symbolized by rows upon rows of binary code. I believe these images, such as [12, 13], reflect the unfamiliarity the individual feels regarding their digital self, a constructed identity that is all at once uncomfortably detailed and frustratingly abstract. Too familiar and yet a stranger, failing to properly account for the complexity of human identity. What seems like anonymity in the form of shadowed figures, whose identifiable characteristics are amiss,

[13] From the article "Big data must become 'people data'" published on March 2014, in the Guardian

may be the public's frightening realization that they are disassembled and turned into data flows. Pieces of code which are then grouped together and turned into algorithms for financial gain or security concerns. Not only is the WEB not anonymous but even when the user voluntarily offers personal information, it obtains far more from them, turning their existence into raw data, ready to be reassembled into something new (Leszczynski, 2015, p.968).

3.4. Gender Construction

Out of the 229 images of my sample, 80 images show people, out of which 74 show discernable figures of people who are not part of a crowd. Within this sample, 47 are men and 17 are women. Overall, 63% of people shown in Big Data images are men, whilst only 37% are women. Though the samples in the Greek newspapers are smaller than the British equivalents, it is interesting to note that *Kathimerini*, a more conservative, center-right newspaper, has no pictures of women whatsoever, while *Eleftherotypia/EFSYN*, a liberal, left-wing newspaper is the only one visualizing a higher number of women. Table 2 shows the percentages of each newspaper in detail:

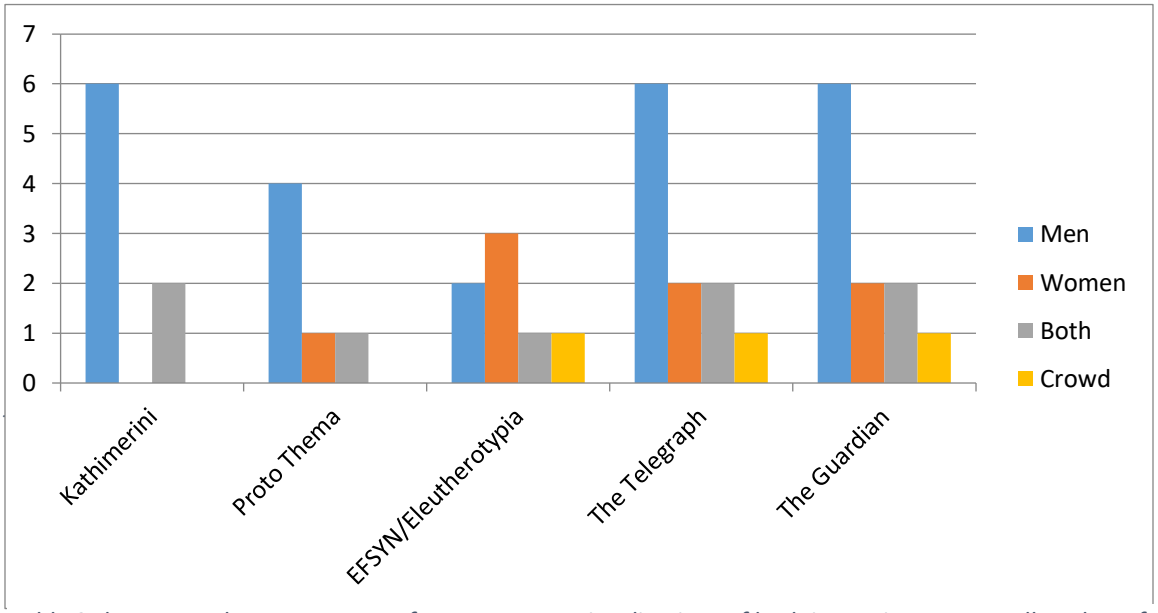


Table 3 showcases the percentages of men, women, visualizations of both in one image, as well as that of crowds, where gender is not easily discernable. In this chart, images where people are absent are not included.

There have been several studies concerned with the construction and representation of gender in computer advertisements, such as the works of Weinstein (1998), Johnson, Rowan and Lynch (2006), Cox (2009) and Tympas, Konsta, Lekkas and Karas (2010), which have offered valuable information about the recurring themes at play.

More specifically, the chapter published by Tympas, Konsta, Lekkas and Karas (2010), entitled "Constructing Gender and Technology in Advertising Images: Feminine and Masculine Computer Parts", interprets images in computer advertisements in a way that analyzes the gender construction taking place within them. Gender bias in computer

advertisements can refer to hierachical relationships formed between men and women, as symbolized in the manner in which they use computers, be it the rigidness of their stance or the job position the content of the work done on the computer itself implies. At the same time, gender is present in parts belonging to the computer, such as the screen and the keyboard, which are constructed as feminine. Female faces are typically shown in computers advertisements "mirrored" on the screen, which ties in to the overall visualization of women typing, since any person actively using the keyboard and looking ahead would have their face placed across the screen and reflected on it (Tympas, Konsta, Lekkas, Karas, 2010, p.203). This work, just as the ones referenced above, strictly deals with computer advertisements and other gadgets belonging to the Third Industrial Revolution. Unlike those papers, my thesis analyzes Big Data which, alongside other emerging technologies such as Robotics, Artificial Intelligence, Nanotechnology and the Internet of Things, belong to the Fourth Industrial Revolution. Nonetheless, I decided to analyze my sample in reference to these past studies to see whether or not gender stereotypes pinpointed in previous researches regarding computers had transferred over to Big Data as well.

The results unquestionably pointed to a continuity between the two technologies and gender construction. I will be presenting a number of themes, including; men and women using computers, particularly the keyboard, and the differences the images exhibit whilst showing the same activity. A secondary aspect of this will be the location, the workplace, and how within it, specific uses of computers signify authority. Additionally, my sample will demonstrate how in some images femininity and masculinity are framed as aspects of the computer itself, an interesting finding previously identified by Tympas, Konsta, Lekkas and Karas (2010). Furthermore, the lack of any attempt at domesticating the technology, as was already noted in a previous chapter, equals a significant absence of the domestic, a stereotypically "feminine" concern. At the same time, stereotypically "masculine" activities, such as sports, are still used to visualize the new technology (Cox, 2009, p.99). The idea of expertise is both present in images of men alone but also in their affiliation with women, in which men adopt the role of the teacher or supervisor. As a final point, throughout the chapter, I will be making note of any exceptions to the cases and themes analyzed in an effort to be thorough.

Research done on an extensive sample of computer ads has shown that whenever men are using a computer, they rarely look at the screen in front of them. The few times men are pictured looking straight at the screen, tables and graphs are depicted on it (Tympas, Konsta, Lekkas, Karas, 2010, p.197). This special context, one that is either financial (i.e. a manager) or scientific (i.e. an engineer), creates the image of the inventor or leader, both positions of prestige and authority. The use of the forefinger tends to be prevalent as well, with the man pointing at key information located on the chart or diagram in question [image 15]. When a keyboard is involved, both hands are usually unoccupied as seen in image [14]. If one of them is placed on the keyboard, an image not uncommon, it is not to type words—a clerical job—but to click on a command, usually signified once again by the use of the forefinger, this time hovering over the Enter key. But not all images visualizing men on computers are located in the workplace. When the impression made is not managerial and the man is not shown in the office, the phone or mouse can be replaced with a number of objects, such as the cup of coffee positioned rather prominently in image [17], which coupled with a laid back posture and the absence of a desk, symbolizes relaxation and casual control of the machinery. Image [16] provides an interesting case where the typical male stance in front of the computer with the telephone at hand is juxtaposed by the female figure next to him, sitting in a similarly unofficial environment, perhaps a café, and yet having an entirely different stance herself. With both hands on the keyboard, eyes sharply focusing on the contents of the screen and her body

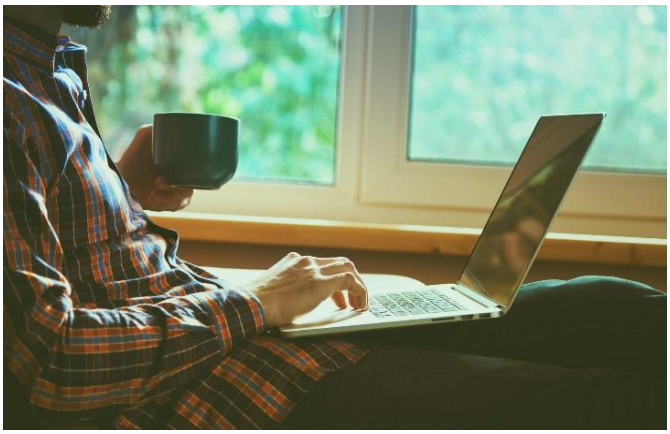
[14] From the article "Beware data overload don't let analysis paralysis stunt your business", published on April '17, in the Guardian

[15] From the article "The first center in Greece for data analysis opportunities" publish on July 2016, in Proto Thema

posture straight and stiff, her bearing signifies routine labour even outside the workplace (Tympas, Konsta, Lekkas, Karas, 2010, p.194).

[16] From the article "Councils call in the geeks to help them solve local problems", published on April 2013, in the Guardian

A generous percentage of my sample exhibits the same traits. True enough, when images portray women on computers, their body language and context diverge from that of men consistently. Women are always shown facing the computer screen directly, using both hands to type on the keyboard, which is coded as feminine (Tympas, Konsta, Lekkas, Karas, 2010, p.205). In the few exceptions where the women are not placing both hands on the keyboard, there is always an additional factor to remind the audience of their femininity, such as in image [18] where the second hand is placed on the woman's abdomen to indicate her pregnancy. Men, on the other hand are consistently shown with one hand holding a phone, a mouse or even a cup coffee while the other rests on the keyboard. The work at hand, be it professional or private, denotes power and decision-making and lacks the characteristics of low-level, routine tasks, such as data-entry.



[17] From the article "How insurers could use 'big data' to charge you more" published on November 2016, in The Telegraph

[18] From the article "Let big data unlock the secrets of our bodies", published on March 2019, in the Guardian

Similar to the computer ads of 1980s and 1990s, a common image is that of the female, low-level

[19] From the article «In December the abolition of paper will begin in public agencies" published on November 2018, in EFSYN

employee, using the computer in the workplace. It is notable that the female workspace tends to lack the freedom and grandeur of the equivalent man's office, the space often pictured as crammed, confined and messy, as can be seen in image [19]. With little space to move, enclosed walls and low, sitting positions, routine work is implied rather overtly, usually that of a secretary, assistant or clerk (Tympas, Konsta, Lekkas, Karas, 2010, pp.195, 197-8).

An additional interesting aspect of the computer screen's coded femininity can be found in images which depict close-ups of women's faces, most specifically, the eyes. In these instances, the female eye substitutes for the full female face, the same way male and female close-ups of hands do in other images. What is mirrored in the eyes is the content of the screen, implying the complete image of the woman staring at it and simultaneously, providing a blatant correlation of the feminine computer screen and the female user herself. Such example can be seen in image [20] (Tympas, Konsta, Lekkas, Karas, 2010, p.203).

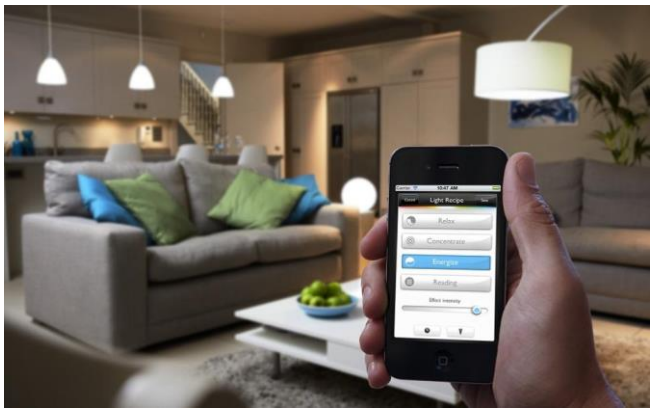
As I mentioned before, the visualization of Big Data lacks the presence of the domestic. And while the domestic and its ties to femininity is gone, masculine-coded activities such as image [21] continue to be represented. This is a continuation of Cox's (2009) findings regarding sports metaphors. One could say that seeing sports portrayed in images visualizing Big Data is hardly shocking, given that one of the major usages for Big Data is in sports analytics. And yet, this does not

[20] From the article "Nobody knows you better than Facebook", published on January 2015, in Kathimerini

adequately explain the absence of the domestic, since another significant area in which Big Data is "flourishing" is smart homes and smart appliances. Domestic appliances, such as washing machines, refrigerators and heaters are undoubtedly feminine-coded and tied to the subject of Big Data and the Internet of Things and yet, I failed to retrieve images of them in my sample, besides a single article that showcased images of a home [image 22] but lacked any people or a sense of domesticity and familiarity. Moreover, the colours of the house in image [22] are cold, predominantly blue and green with a distinct lack of red and its warm hues, unlike the images of the domestic I encountered in an earlier chapter, regarding the computer and the *home,* back in the 1970s and 1980s.

[21] From the article "Hole in one How Big Data can turn players into winners", published on September 2014, in the Guardian

An additional analysis could posit that visualizing Big Data with images such as [21] is a deliberate decision, reminiscent of past advertisements for computers showing men in open spaces, close to nature, in the midst of entertaining activities, such as sailing (Tympas, Konsta, Lekkas, Karas, 2010, p.194). This kind of visualization of technology, similarly presented here, effectively removes "machinery" and all the threatening or negative connotation it might entail with it and instead shows man in an open space, controlling and using the technology but doing it in a way that seems natural and henceforth, approachable and more fun.



[22] From the article "LG: technological trends for 2015". Published on December 2014, in Kathimerini.

[23] From the article "Separating the truth from the buzz in social media", published on September 2015, in the Guardian

Another interesting theme appears in images where men and women use a computer in tandem, such as in image [23]. In images such as this, a relationship of power forms between them, where the man traditionally takes the role of the teacher or manager and the woman takes the role of the student or employee. In the words of Johnson, Rowan and Lynch (2006), "expertise [in the computer field] is associated with technological confidence, independence, mastery, and authority" (p.7). When images consistently place women in passive positions, sitting and following instructions from men, whose standing position and explanatory role associates them with power, knowledge and expertise, technological skill is connected to masculinity.

An exception to the passive placement of women regarding computers (and ultimately Big Data) seems to arise in image [24], where a notable person is portrayed instead of a nameless male or female. The woman in question is a distinguished figure and the article refers to her specifically by name. Furthermore, her posture and placement within the image does not follow stereotypical feminine traits of advertisements past. As was mentioned earlier in this thesis, this could partly be because of the use of stock images. Because the already existing sample of options when choosing a generic picture to accompany an article, contains stereotypical placement of male and female figures, the chances of the selected picture having gender bias is statistically high. And yet, approaching it with a combination of colour theory and gender commentary, as already presented in this thesis, this exception evidently becomes not an exception at all but an

[24] From the article "Genevieve Bell: 'Humanity's greatest fear is about being irrelevant", published on November 2016, in the Guardian

attempt at domesticating Big Data, using methods not unlike those of the past. The image boasts warm hues, predominantly orange and yellow, with no blue whatsoever. This is a noteworthy absence, given the overwhelming blueness present in most Big Data images. The woman pictured here, Genevieve Bell, although named and referenced extensively in the article, is in fact an Australian anthropologist. So, even though she discusses Big Data, she is not a figure possessing traditional know-how and expertise regarding the technology, nor does she belong to a technological or industrial profession. Her humanities background and the warm colours surrounding her are meant to domesticate the technology in a way that offers remarkable continuity, spanning from the third to the fourth industrial revolution, a link that was in fact systematically and repeatedly strengthened during the analysis of my sample in this chapter.

Similarly to the case of Dataveillance, one could raise the question, why are these interpretations or visualizations of the female identity problematic? Why is the construction of gender in both the third and the fourth industrial revolution treated as a negative, as an issue in this thesis? Surely enough, a considerable portion of those who read these articles and see these images do not necessarily understand the biases being presented, nor do they find them problematic in any way. And yet, the internalization of stereotypical behaviours, uneven power relationships and biased codification of feminine and masculine traits presents a serious, social problem, analyzed by many scholars (Bevan & Learmonth, 2013; Yansen & Zukerfeld, 2014; Hong, 2016). Part of why it is so pervasive is the fact that it is not obvious. The images analyzed in this chapter are not overtly sexist, nor do they clearly denote negative stereotypes, such as the oversexualization of the female body present in some computer advertisements of the past (Tympas, Konsta, Lekkas, Karas, 2010, p.199) and yet, they provide an interesting case of how implicit bias works in journalism.

3.5. Analogue versus Digital

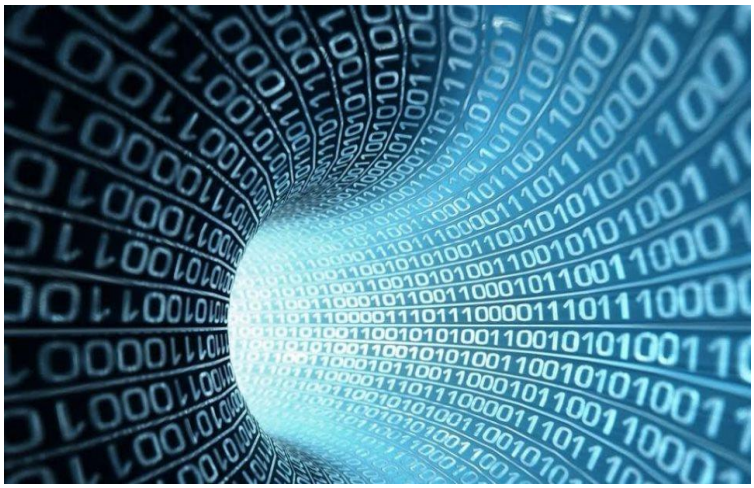
Big Data is habitually presented as having to do with digital technology, non-analogue technology. As, however, several authors in a recent volume on the history of Big Data have argued, Big Data depends on the laborious transformation of analogue media into digital ones (Strasser, Edwards, 2017; Aronova, von Oertzen, Sepkoski, 2017). The importance of retrieving the persistence of the analogue within what is presented as digital has also been argued by A. Tympas (2018). Furthermore, the importance of studying the digital-analogue relationship in computing advertisements has been argued by Tympas et al., who have shown that it is connected to labour and gender relationships (Tympas, Konsta, Lekkas, Karas, 2010). This is why I now turn my attention to the digital-analogue relationship as displayed in images about Big Data as a technology independent of the labour needed to produce a computing analogy. In an earlier chapter, I mentioned the rich history of Big Data, dating back to the pre-electronic era. As Aronova, von Oertzen and Sepkoski (2017) describe, the argument in favour of classifying Big Data as digital is the way they are stored, translated into pieces of binary code, uploaded to the Cloud and housed in massive digital databases. This procedure is not in any way concealed from the public eye, even if the average netizen is not completely aware of how the process works or how much physical space it takes to stow digital data. Due to the extensive publicity of this aspect of Big Data, a noteworthy percentage of images used to visualize the emerging technology, has to do with binary coding. In all five newspapers I based my research on, there is a plethora of images accompanying Big Data articles with screens filled with rows upon rows of ones and zeroes, mostly placed on a blue background to fit the overall theme of blue oversaturation present in Big Data imagery [image 25], but sometimes also in a black background with the numbers themselves coloured luminous green or orange.

[25] From the article "Public sector needs to think big when it comes to data" published on the Guardian, in April 2012

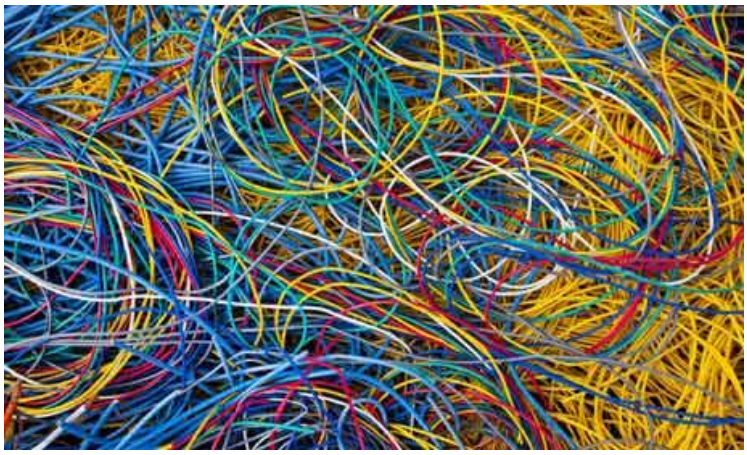
Despite this, there are several instances in my sample where Big Data is visualized with cables and other physical entities, usually belonging to immense computing systems. This is interesting in two respects. The first, is the analogue nature of cables as a material, whereas the second, is the allusion to the early analogue history of computers, as showcased in image [26], providing a link between the third industrial revolution, also known as the digital revolution, and the fourth industrial revolution where Big Data is located, alongside other emergent technologies such as Artificial Intelligence and Robotics.

[26] From the article "How big data is transforming public services – expert views", published in the Guardian, in April 2014

This dual visualization of Big Data as either purely digital or made of analogue components, conceivably highlights an intrinsic confusion regarding the nature of Big Data itself. Symbolizing Big Data solely as binary code and electronic signals disregards its vast history, tracing as far back as the Library of Alexandria, which arguably had the largest collection of manuscripts during antiquity and provides a fitting example of an analogue case for "big data". Concurrently, symbolizing Big Data using either imagery dating back to the third industrial revolution or instruments that are commonly coded as analogue, provides a misinformed and outdated view on how the technology currently works.



[27] From the article "Big Data against climate change", published on Proto Thema, in December 2014



[28] From the article "The big data race reaches the City", published in Telegraph, on October 2016

An additional important insight from the dual nature of Big Data, as they are depicted in these images, is that this dyad is not strictly divided by chronological "borders". It would be an oversight to say that Big Data are analogue only when they belong to the past, whereas now they are digital, vast amounts of information continuously uploaded to the "cloud". The idea of the cloud, even its very name, gives a sense of weightlessness, a metaphorical cloud storing countless bytes of data but not "taking any space" for it is not physical, it is *digital*, just ones and zeroes floating in the WEB. This could not be further away from the truth. For every additional terabyte, exabyte or even vottabyte of data stored in the cloud, the infrastructures necessary to contain it are not that different from the Library of Alexandria. In the same way that a progressively larger collection of books needs more space to house it, Big Data's massive datasets need larger structures to hold them inside, stowed away in enormous computers, handled by personnel responsible for upkeep and protection of the innumerable amounts of information. These buildings have both financial and geographical needs, money for maintenance and land for expansion.

[29] From the article "Metadata: The rise of data on the web and in search", published in The Guardian, in May 2013

The images depicting Big Data as binary code far outweigh those depicting them in a more physical or analogue form. And that is partially responsible for the promotion of the idea that Big Data is something enigmatic, shadowy and hard to identify, alienating in its vagueness. An idea that spreads

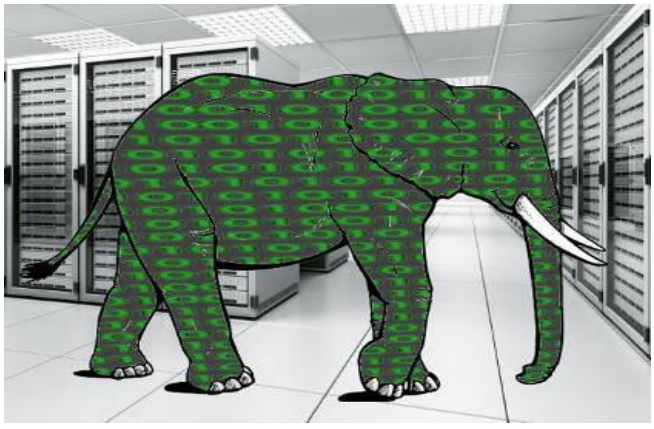
into what in earlier chapters was described as Big Data-related Anxiety. Electric signals transferring pieces of code to some insubstantial place, although untrue, is what rows of binary code floating in solid colour with no structure, weight or other elements in the background signify. And in that sense, not only Big Data's analogue nature is concealed, but digital also becomes equivalent to intangible, perhaps even uncontainable and most certainly out of the realm of human agency.

[30] From the article "A development data revolution needs to go beyond the geeks and bean-counters", published in the Guardian, in October 2013

3.6. Liberal vs Conservative, Greece vs UK:

In this chapter, I will be providing some quantitative insight on my sample in correspondence to the political beliefs of the newspapers in question. Then, by using the quantitative results that I have gathered thus far, I will be comparing Greece and the U.K. to determine the similarities or divergences in the way they visualize Big Data in their newspapers. My primary aim is to calculate the amount of negative imagery existing in each newspaper and to figure out whether or not the percentages have any correlation with how liberal or conservative each paper is. My original assumption, when beginning this thesis, was that there is in fact a correlation between negative Big Data visualizations and liberalism. My results, which I will provide below, confirm my assumption for the better part. In the following text, I will be commenting on this correlation and also providing a possible explanation regarding any exceptions.

I will begin with the newspapers which are commonly identified as conservative. Approximately 13% of the *Kathimerini* sample has negative connotations, such as pictures of surveillance, hackers and privacy and security concerns. Due to being a conservative newspaper, it is no surprise seeing it downplay the negative aspects of Big Data but having such small percentage is nonetheless noteworthy. In turn, The Telegraph's results are more balanced. A 29% of *The Telegraph* has negative connotations, mostly in the form of shadows of people, surveillance of crowds and drones, who are somewhat neutral in appearance but nonetheless can and do imply dataveillance.



[31] From the article "The modern «Big Brother», published on May 2013, in Eleftherotypia.

Meanwhile, liberal or populist newspapers, such as *Proto Thema, Eleftherotypia/EFSYN* and *The Guardian* have significantly higher percentages. More specifically, 23% of *Proto Thema* has negative connotations, such as emotional manipulation due to social media platforms, breaches of privacy and hacking. Similarly, 36% of the *Eleftherotypia/EFSYN* sample has negative connotations,

more overt than those in the other newspapers, showcasing Big Brother references [image 31], privacy breaches and data manipulation via platforms such as Facebook and the prevalence of Fake News. *The Guardian*'s percentage is approximately 34% of the images having negative connotations in the form of surveillance, hackers, privacy and security concerns. What is noteworthy, is that when negative imagery in the U.K. press is further split into two categories; Dataveillance and issues of privacy in the WEB, the percentages on each newspaper change. That will be apparent in the graphs situated further below. For now, although dealing with a larger sample, the percentage of negative imagery regardless of specific themes is the same in *The Guardian*, as it is in The Telegraph, a conservative paper whose interests lie in the protection of the legitimacy of the government, corporate companies and organizations. A potential explanation for *The Guardian*'s fairly small percentage is the newspaper's preference for academic sources when writing articles concerning science.

The chart below showcases in percentages the frequency in which each motif appears in all three Greek newspapers present in my sample. It is interesting to note that Kathimerini, coloured blue, exhibits the higher percentage of gender construction out of all three newspapers while at the same time, it exhibits the lowest percentages of Big Data images related to Dataveillance and issues of privacy and misconstrued identities online.

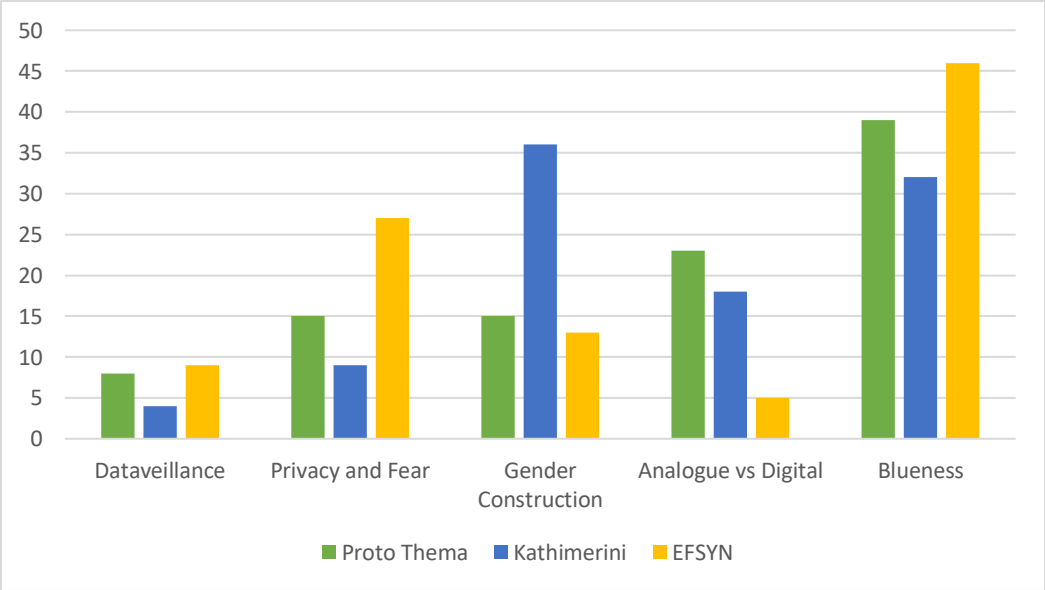


Table 4 showcases the percentages of the chief five themes present in my empirical sample of Big Data imagery, as calculated using only the three Greek newspapers, Proto Thema (center/far right, liberal), Kathimerini (right, conservative) and EFSYN/Eleftherotypia (left, liberal).

Regarding *EFSYN/Eleftherotypia*, a liberal newspaper, the percentages of gender construction are much lower, in fact the lowest of the three. Additionally, it reports the highest amount of imagery related to both dataveillance and privacy issues online, with the latter significantly higher than the other two newspapers. Similarly, *Proto Thema*, a newspaper that is commonly labeled right-wing showcases percentages closer to *EFSYN*, rather than *Kathimerini*, which can perhaps be attributed to its populist agenda. *EFSYN*'s notably low percentage of Big Data visualized in the form of either digital or analogue components is another significant percentage differentiation between the three newspapers. Possibly a result of the higher visualization of privacy concerns, images and articles that most definitely incite journalistic sensationalism over the less enthusiastic and less mainstream imagery of machinery.

The chart below showcases in percentages the frequency in which each motif appears in the two British newspapers present in my sample. Here, the differentiation between the two papers is more balanced, the difference in percentages less apparent, with the exception of the colour blue and the images related to Dataveillance.

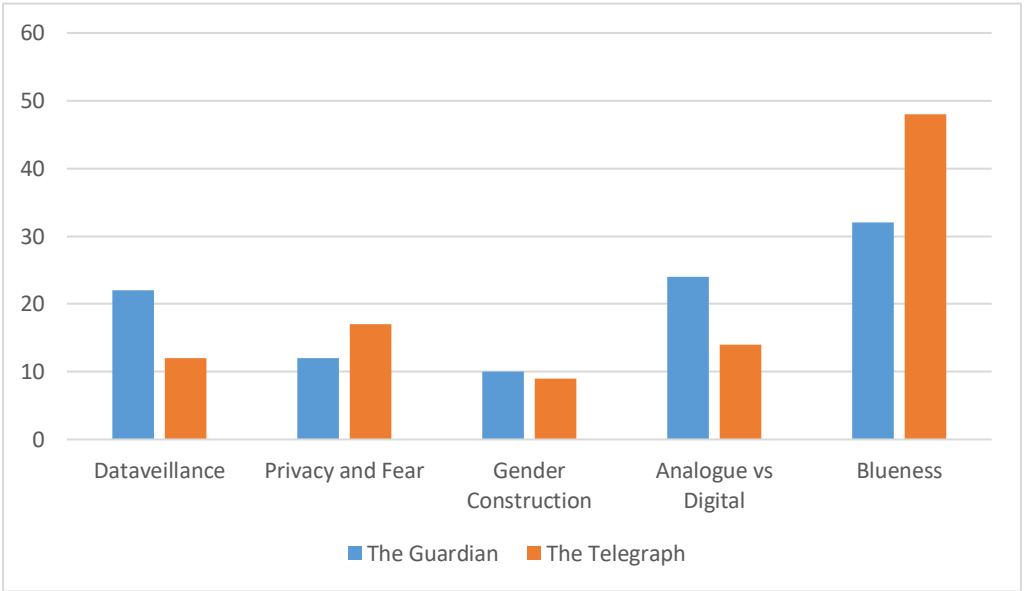


Table 5 showcases the percentages of the chief five themes present in my empirical sample of Big Data imagery, as calculated using only the two British newspapers, The Guardian (left, liberal), and The Telegraph (right, conservative).

Beginning with the latter, *The Guardian* visualizes Big Data in the form of Dataveillance far more frequently than The Telegraph, but interestingly enough, when it comes to issues of privacy online and the fear of misconstrued digital identities, *The*

Telegraph produces more related content. Regarding gender construction, the percentages are closely matched, with *The Guardian* exhibiting the subject matter 1% higher than *The Telegraph*. Finally, imagery of Big Data as either analogue or digital is more present in *The Guardian* than *The Telegraph*, possibly due to the aforementioned tendency of *The Guardian* towards academically-inclined journalism. *The Guardian*'s significantly higher numbers in Dataveillance imagery can also be attributed to liberal newspapers' tendency to pursue anti-governmental issues more than their conservative counterparts.

Due to that, I can safely say there is a correlation between political beliefs and negative visualizations of Big Data. Conservativism is concerned with the perseverance of institutions such as religion and government and is a major proponent of tradition, wishing for the things to continue "as they were" (Heywood, 2012, p.68). On the other hand, Liberalism is concerned with liberty, human rights, gender and racial equality and the active limitation of governmental power and authority (Dune, 1993). Although Big Data is an emerging technology and it would seem sensible that a conservative newspaper interested in upholding traditions would react adversely to it, the numbers say something different. The reasoning behind this decision lies in the active role of the government in dataveillance. Dataveillance and incessant data-gathering are undoubtedly the most widely visualized negatives of Big Data in my sample but due to the involvement of the state, such as in the case of Snowden in 2013, pro-government, conservative papers consciously discuss them less than their liberal counterparts.

Finally, at the introduction of this thesis, I decided to include newspapers from the U.K. as well as Greece, to ascertain whether or not Greece presented a unique case in the visualization of Big Data, an idiosyncratic example which does not affect the academic community outside the strict borders of the country. The chart below [table 6] presents a comparison of the two countries. For this comparison to be possible, I have placed all my findings in only two categories, British newspapers and Greek newspapers, essentially removing the political affiliation of the newspapers as a factor of comparison. The numerical results, showcased in percentages and not numbers of Big Data images in each country, provide proof that Greece is not a unique case. At the same time, the results reveal some substantial dissimilarities between the two countries. The use of the colour blue to

visualize Big Data is equally frequent in both countries and the difference in percentage is small enough that a comparison is essentially insignificant.

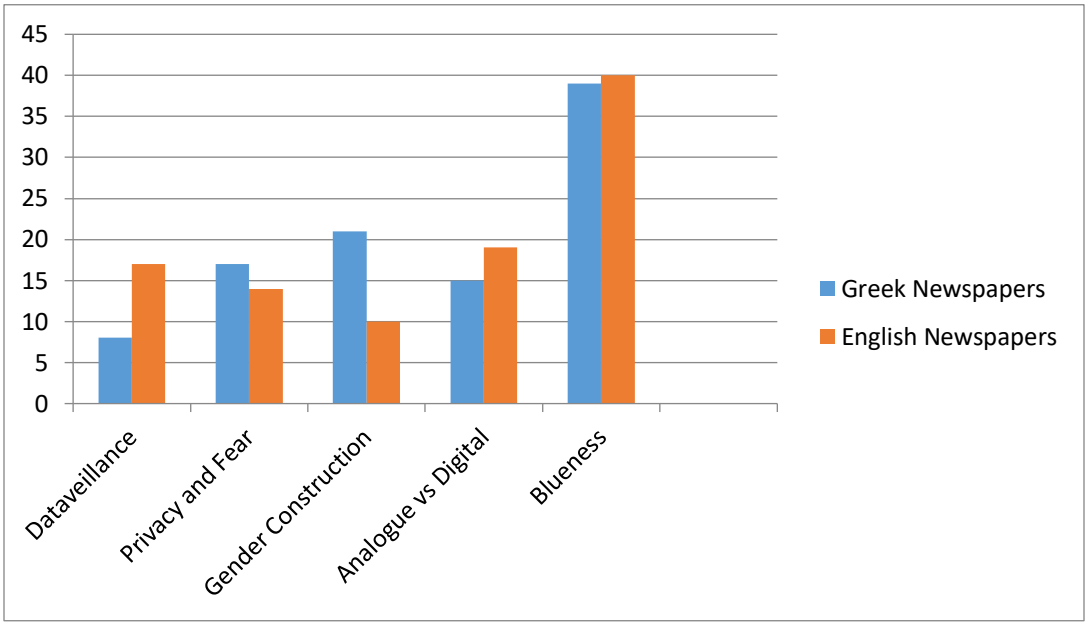


Table 6 showcases the percentages of the chief five themes present in my empirical sample of Big Data imagery, as calculated using the sum of all five newspapers, split into the two countries involved in the empirical research, Greece and the U.K.

The first noteworthy percentage difference comes in the visualization of Big Data as either digital or analogue, where British newspapers hold a higher frequency rate. In a similar percentage differentiation, Greek newspapers evidently visualize issues of privacy on the WEB, especially in the form of Fake News and Facebook data-gathering more than the British newspapers. Moving on to themes where the percentage gap is much wider, Greek newspapers score more than 10% higher in Gender Construction, whereas British newspapers score 10% higher in Dataveillance imagery. The two most significant percentage differences are Dataveillance and Gender Construction, where the latter's percentage gap is the widest in the entire comparison. A possible attribution to the lower frequency of Dataveillance imagery in Greek newspapers could be that the Snowden incident was less prevalent in news in Greece at the time, rather than in the U.K. The higher frequency of Gender Construction imagery could be attributed to higher percentages of gender stereotyping and gender bias present in Greek society. But these are only personal hypotheses. What is the most important outcome of this comparison is that Greece is not

an idiosyncratic case and can be compared to other countries, both within Europe and outside of it, both in this thesis and also in future research pursued by other scholars.

Conclusion and suggestions for further research

This research aimed to identify repeating motifs existing within the images used in articles to visualize the emerging technology that is Big Data. To achieve this, a sample was created by tracing images accompanying articles referring to Big Data, retrieved from three Greek newspapers and two British newspapers. The study has asked the following questions:

(1) Do images visualizing Big Data evoke predominantly negative emotions to the public, regardless of the content of the article they accompany?
(2) Do images visualizing Big Data contain unconscious gender bias, similar to advertisements of computers belonging to the Third Industrial Revolution?
(3) Do liberal newspapers use negative imagery to communicate Big Data more than their conservative counterparts?

Based on a mostly qualitative and sometimes quantitative analysis of my empirical sample, my research brought forth five chief themes, blueness, dataveillance, fear regarding loss of privacy and misconstrued identities in the WEB, construction of gender and attempts to visualize the technology itself either in a digital or an analogue form. By first assessing the public's response to issues such as dataveillance and privacy on the WEB using secondary literature, it was possible to highlight the ways in which such imagery reflects negative and stressful emotions irrevocably linked to these activities. In answer to the primary research hypothesis, quantitatively, the conjoined percentages of imagery related to dataveillance and issues of privacy and identity in the WEB, account for 28% of the total sample belonging to all five newspapers. Blueness accounts for 39.5% of the total

sample but is excluded in this quantitative analysis, given the fact that it is an overwhelming trait most images possess, usually secondary to another motif, such as visualizations of coding, women in the workplace or security cameras. With the colour blue placed outside my calculations, imagery of Big Data that evokes predominantly negative emotions is indeed the most frequently present in my sample and thus, my primary research hypothesis is confirmed.

To answer the second research hypothesis, I began by tracing all images within my sample showing men and women, providing a quantitative analysis and consequently, analyzing their content qualitatively, in order to highlight implicit gender bias present in the visualizations. Juxtaposing my findings with work done in the past on computer advertisements belonging to the Third Industrial Revolution, I managed to pinpoint the various ways in which gender stereotypes transferred through the decades, from computers to Big Data and while their identification proved more challenging, their presence had nonetheless persisted. Besides confirming my research hypothesis, my analysis on this part of the sample, allowed me to ascertain the deceptive ways in which overtly sexualized images of women were replaced with less explicit, but no less problematic representations of stereotypical femininity.

The results of the third and final research hypothesis both confirmed and disproved the initial question. While *The Telegraph* and *Kathimerini*, both conservative and right-wing newspapers, exhibited low percentages of dataveillance, *The Telegraph* visualized concerns regarding safety in the WEB more than its liberal counterpart, *The Guardian*. *Kathimerini*'s findings stayed on par with the research hypothesis, providing the highest confirmation rate to the original hypothesis but the least amount of interest in this

idiosyncratic case. These findings indicate that dataveillance, a subject closer tied to governmental participation, was deliberately ignored by conservative newspapers, whereas issues regarding the internet were more freely publicized. Hence, the original hypothesis was not nuanced enough to properly divide negatively-coded Big Data images into the two parts necessary to explain the conflicting empirical results. A secondary finding that partly disproves the original research hypothesis and introduces a new element, whose further exploration could prove of interest, is the sample retrieved from the newspaper *Proto Thema*. *Proto Thema* is a right-wing newspaper, which unlike *Kathimerini* and *The Telegraph*, is also populist. The percentages in negatively-coded imagery are closer to those of *EFSYN/Eleftherotypia*, a liberal, left-wing newspaper than the right-wing, conservative *Kathimerini*. This kind of discrepancy implies that perhaps it is not only liberalism that leads newspapers in producing more negative Big Data imagery but populism also. *Proto Thema* is not liberal and yet produces similar content to liberal newspapers. Conceivably, an answer to the research hypothesis is that both liberalism and populism play a role in the kind of visualization of Big Data and newspapers who are right-wing or conservative can still produce a noteworthy amount of negative media concerning Big Data, assuming they are also populist. Furthermore, conservative newspapers certainly downplay the negative aspect of Big Data but mostly when that has to do with state involvement. When private companies are involved in a more apparent manner and any government involvement is obscure, conservative newspapers oftentimes produce negative Big Data content in larger amounts than their liberal counterparts.

A limitation to this study has been the deliberate decision to exclude any analysis regarding the choice of imagery on behalf of the journalists. Instead, I have chosen to focus

only on the end result, the published article, without providing any justification or reasoning behind the decision to publish each specific photo. While abstaining from analyzing each part of the procedure necessary to publish these images somewhat limits the final results, this approach has allowed me to analyze in greater depth what comes after, providing new and interesting insights into the communication of Big Data to the public and vice versa, that would not be possible, had I chosen to divide both my time and length of the thesis in handling both parts simultaneously.

To better understand the implications of these results, future studies pursued by scholars could address issues related to geography, content and even types of media. Further research is needed to determine whether or not other countries in Europe, as well as countries situated outside of the European borders, produce the same kind of findings or differ significantly. Such extension juxtaposed to my own empirical results would offer a broader and fuller understanding of the visualization of Big Data in newspapers and how said visualization affects the public and vice versa. A secondary aspect in which future research would produce valuable and interesting results is the way the Fourth Industrial revolution is visualized in media in general. Big Data is an important part of the Fourth Industrial Revolution but other emerging technologies belong to the same technological wave that so much affects society as a whole. In that sense, scholars and researchers could produce a similar empirical work to mine, a content analysis-approach to the visualization of the other chief technologies belonging to the Fourth Industrial Revolution, such as Artificial Intelligence, Robotics, Nanotechnology, Biotechnology and Fully Autonomous Vehicles. It would be interesting to see patterns surfacing within them that are either unique to the visualization of each technology or perhaps offer similarities to others studied before

them. Finally, the types of media used to visualize these technologies should not be ignored, nor their potential effect on the visualization itself. My own work deals with the visualization of Big Data in printed media, as showcased in their digital counterparts, but other future research can focus on different types of media, such as television and film.

References

Abed, Y., & Chavan, M. (2019). The Challenges of Institutional Distance: Data Privacy Issues in Cloud Computing. *Science, Technology and Society*, *24*(1), 161–181. https://doi.org/10.1177/0971721818806088

Allen, W. (2019). *The Politics of Big Data, Migration, and Mobility*. Presentation, Greece. Further information: http://www.ekt.gr/el/news/23216

Andrejevic, M. & Gates, K. (2014). Big Data Surveillance: Introduction. *Surveillance and Society. 12*. 185-196. https://doi.org/10.24908/ss.v12i2.5242

Aronova, E., von Oertzen, C., & Sepkoski, D. (2017). Introduction: Historicizing Big Data, *Osiris, 32*, 1-17, https://doi.org/10.1086/693399

Aslam, M. M. (2006) Are You Selling The Right Colour? A Cross-cultural Review of Colour As A Marketing Cue, *Journal of Marketing Communications, 12(1)*, 15-30, https://doi.org/10.1080/13527260500247827

Benford, R. D. & Snow, D. A. (2000) Framing Processes And Social Movements: An Overview And Assessment. *Annual Review Of Sociology, 26, (pp. 611-639)*, Annual Reviews.

Bevan, V., & Learmonth, M. (2013). 'I wouldn't say it's sexism, except that … It's all these little subtle things': Healthcare scientists' accounts of gender in healthcare science laboratories. Social Studies of Science, 43(1), 136–158. https://doi.org/10.1177/0306312712460606

Bijker, W. E. (Olsen, J. K. B., Pedersen, S. A. & Hendricks, V. F.) (2009). *Social Construction of Technology*. Wiley-Blackwell. https://doi.org/10.1002/9781444310795.ch15

Bultitude, K. (2011), The Why and How of Science Communication. In: Rosulek, P., ed. "Science Communication". Pilsen: European Commission. https://www.scifode-foundation.org/attachments/article/38/Karen_Bultitude_-_Science_Communication_Why_and_How.pdf

Cadwalladr, C. & Graham-Harrison, E. (2018, March 17). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. The Guardian. Retrieved from https://www.theguardian.com

Collins, H. M. & Yearly, C. (1992) Epistemological Chicken. In Andrew Pickering, *Science as Practice and Culture*. University of Chicago Press.

Colman, M. A. (2008) A Dictionary of Psychology (3 ed.). Oxford, U.K.: Oxford University Press.

Cox, A., (2009). Visual Representations Of Gender And Computing In Consumer And Professional Magazines. *New Technology, Work and Employment, 24(1),* 89-106, http://dx.doi.org/10.1111/j.1468-005X.2008.00220.x

Crawford, K. & Schultz, J. (2014). Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms, *55 B.C.L. Rev. 93,* 93-128, http://lawdigitalcommons.bc.edu/bclr/vol55/iss1/4

Cressman, Darryl. (2019). A Brief Overview of Actor-Network Theory: Punctualization, Heterogeneous Engineering & Translation. http://summit.sfu.ca/item/13593

Degli-Esposti, S. (2014). When Big Data Meets Dataveillance: The Hidden Side Of Analytics. *Surveillance & Society. 12.* 209-225. https://www.researchgate.net/publication/262493771_When_big_data_meets_dataveillance_The_hidden_side_of_analytics

Dourish, P., & Gómez Cruz, E. (2018). Datafication and data fiction: Narrating data and narrating with data. Big Data & Society. https://doi.org/10.1177/2053951718784083

Dunn, J. (1993) Western Political Theory in the Face of the Future (2nd ed.). Cambridge, U.K.: Cambridge University Press

Frosh, P. (2001). Inside The Image Factory: Stock Photography And Cultural Production. *Media, Culture & Society*, *23*(5), 625–646. https://doi.org/10.1177/016344301023005005

Goldsmith, O. (1764). The Traveller. Retrieved, May 23, 2019, from https://readytogoebooks.com/OG-Traveler.html

Helbing, D. (2015). Thinking Ahead - Essays on Big Data, Digital Revolution, and Participatory Market Society. Zürich, Switzerland: Springer.

Heywood, A. (2012). *Political Ideologies: An Introduction*. (5) Palgrave Macmillan

Hildebrandt, M. (2013) Slaves to Big Data. Or Are We? *Derecho y Política*, *17*, 7-44, https://www.raco.cat/index.php/IDP/article/viewFile/303366/393038

Hirsch, R. (2004) Exploring Colour Photography: A Complete Guide. London, U.K.: Laurence King Pub.

Hong, R. (2016). Soft skills and hard numbers: Gender discourse in human resources. Big Data & Society. https://doi.org/10.1177/2053951716674237

Hsieh, H.-F. & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis, Qualitative Health Research 15(9), 1277-88 https://doi.org/10.1177/1049732305276687

Hume, D., Norton, D. F., & Norton, M. J. (2000). A treatise of human nature. Oxford: Oxford University Press.

Johnson, N. F., Rowan, L. & Lynch, J. (2006). Constructions Of Gender In Computer Magazine Advertisements: Confronting The Literature. *SIMILE Studies In Media & Information Literacy Education, 6(1),* 1-11, https://ro.uow.edu.au/edupapers/19/

Labrecque, L. & Milne, G. (2011). Exciting Red And Competent Blue: The Importance Of Color in Marketing. *Journal of The Academy of Marketing Science. 40(5),* 711-727, https://doi.org/10.1007117470100245

Latour, B. (2005). *Reassembling the Social An Introduction to Actor-Network-Theory*. Oxford University Press

Leszczynski, A. (2015). Spatial big data and anxieties of control. *Environment and Planning D: Society and Space, 33*(6), 965–984. https://doi.org/10.1177/0263775815595814

Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, Critique. *Big Data & Society,* 1-13 https://doi.org/10.1177/2053951714541861

Maier, M., Rothmund, T., Retzbach, A., Otto, L. & Besley, J. C. (2014) Informal Learning Through Science Media Usage, *Educational Psychologist, 49:2*, 86-103, https://doi.org/10.1080/00461520.2014.916215

Mayer-Schönberger, V. (2014) Big Data: A Revolution That Will Transform How We Live, Work, and Think. Boston, Massachusetts: Eamon Dolan/Mariner Books

Michael, M., & Lupton, D. (2016). Toward a manifesto for the 'public understanding of big data.' *Public Understanding of Science, 25*(1), 104-116. https://doi.org/10.1177/0963662515609005

Office for National Statistics. (2013). *Internet access - households and individuals, Great Britain: 2013.* Retrieved from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2013-08-08

Office for National Statistics. (2015). *Internet access - households and individuals, Great Britain: 2013.* Retrieved from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2015-08-06

Office for National Statistics. (2017). *Internet access - households and individuals, Great Britain: 2013.* Retrieved from: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2017

Pan, Z., & Kosicki, G. M. (1993). Framing Analysis: An Approach To News Discourse, *Political Communication, 10(1)*, 55-75, https://doi.org/10.1080/10584609.1993.9962963

Papacharissi, Z. (2015). The unbearable lightness of information and the impossible gravitas of knowledge: Big Data and the makings of a digital orality. Media, Culture & Society, 37(7), 1095–1100. https://doi.org/10.1177/0163443715594103

Pentzold, C., Brantner, C., & Fölsche, L. (2019). Imagining Big Data: Illustrations Of "Big Data" In US News Articles, 2010–2016. *New Media & Society, 21(1),* 139–167. https://doi.org/10.1177/1461444818791326

Pinch, T. J., & Bijker, W. E. (1984). The Social Construction of Facts and Artefacts: or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, *14*(3), 399–441. https://doi.org/10.1177/030631284014003004

Pink, S., Lanzeni, D., & Horst, H. (2018). Data anxieties: Finding trust in everyday digital mess. *Big Data & Society*, 1-14, https://doi.org/10.1177/2053951718756685

Reed, L. (2000). Domesticating The Personal Computer: The Mainstreaming Of A New Technology And The Cultural Management Of A Widespread Technophobia, 1964-, *Critical Studies in Media Communication, 17(2)*, 159-185 https://doi.org/10.1080/15295030009388388

Short, D. (2013). The Public Understanding Of Science: 30 Years Of The Bodmer Report. *School Science Review, 95*, 39-44. https://www.researchgate.net/publication/255712425_The_public_understanding_of_science_30_years_of_the_Bodmer_report

Silverstone, R. & Hirsch, E. (1993). Consuming Technologies: Media and Information In Domestic Spaces. *Man, New Series, 28(3)*, 614-615 https://doi.org/10.2307/2804260

Sismondo, S. (2008). Science And Technology Studies And An Engaged Program. *The MIT Press*, pp.13-31, https://www.researchgate.net/publication/261948467_Science_and_Technology_Studies_and_an_Engaged_Program

Smith, G. J. (2018). Data doxa: The affective consequences of data practices. *Big Data & Society*. https://doi.org/10.1177/2053951717751551

Sonesson, G. (1994). Pictorial Semiotics. Gestalt Theory, and the Ecology of Perception. https://doi.org/10.1515/9783110227789.0345

Strasser, B. J., Edwards, P. N. (2017). Big Data Is the Answer … But What Is the Question? *Osiris, 32*, 328-345. https://doi.org/10.1086/694223

Sumner, J. (2012). Today, Computers Should Interest Everybody. The Meanings of Microcomputers, *Studies in Contemporary History*, *Online-Ausgabe, 9, H. 2*, 307-315 http://www.zeithistorische-forschungen.de/2-2012/id=4725

Tympas, A. & Konsta, H., Lekkas, T. & Karas, S. (2010). Constructing Gender And Technology In Advertising Images - Feminine and Masculine Computer Parts. In T. J. M. Misia. *Gender Codes: Why Women Are Leaving Computing*, (pp.187 – 209, Ch.9). Wiley-IEEE Computer Society Press.

Τύμπας, Α. (2018). ΑΝΑΛΟΓΙΚΗ ΕΡΓΑΣΙΑ, ΨΗΦΙΑΚΟ ΚΕΦΑΛΑΙΟ. Αθήνα, Ελλάδα. Angelus Novus.

Van Dijck, J. (2014). Datafication, Dataism and Dataveillance: Big Data Between Scientific Paradigm and Ideology. *Surveillance & Society 12(2)*, 197-208 https://doi.org/10/24908/1224776

Weigold, M. F. (2001). Communicating Science: A Review of the Literature. *Science Communication, 23*(2), 164–193. https://doi.org/10.1177/1075547001023002005

Wynne, B. (1995). 17 Public Understanding Of Science. In Jasanoff, S., Markle, G. E., Peterson, J. C. & Pinch, T. *Handbook of Science and Technology Studies*, 361-389. Thousand Oaks. SAGE Publications, Inc. http://dx.doi.org/10.4135/9781412990127.d24

Yansen, G., & Zukerfeld, M. (2014). Why Don't Women Program? Exploring Links between Gender, Technology and Software. Science, Technology and Society, 19(3), 305–329. https://doi.org/10.1177/0971721814548111