The Human Infrastructure

of

Artificial Intelligence

Master Thesis by Carlo Mervich

EXAMINATION COMMITTEE

First Supervisor: Dr. A.Weber
Second Supervisor: Prof. Dr. Ir. M. Boon

University of Twente, Enschede, the Netherlands
Faculty of Behavioural, Management and Social Sciences (BMS)
MSc Philosophy of Science, Technology and Society (PSTS)
17th of August 2020

Abstract

The contemporary data-driven paradigm of Artificial Intelligence (AI) envisions an automated future, where human functions will be less and less relevant for society due to the ubiquity of AI-based technologies. Such expectations are however contradicted by the large amount of invisible human labour involved in the process of data labelling and curation required today to develop and train machine-learning models which lie at the heart of AI technologies. Amazon Mechanical Turk, Scale-AI and other “human-in-the-loop” systems involve many humans to teach artificially intelligent technologies how to recognize objects, answer questions about the weather, or drive autonomously through the streets. In order to identify, explore and map the role of such a network of people which I define as *human infrastructure*, this thesis draws upon scholarship in the field of Critical Infrastructure Studies (STS). Building on the concept of “infrastructural inversion”, my thesis analyses the development of AI through the lens of the human infrastructure that underlies it. By doing so, it first identifies the mechanisms that make workers invisible. Second, it discusses ethical concerns with respect to workers’ labour conditions. Third, it highlights epistemological issues related to data processing. As a last step, it analyses how the involvement of humans actually shapes the development of AI systems. By adopting a human-centered approach, this thesis provides a critical view on many present-day conceptualizations of AI.

Table of Contents

List of figures

Introduction

For more than sixty years, computer scientists, engineers, linguists, philosophers and scholars have been studying and working on Artificial Intelligence (AI) systems that aim at replicating human intelligence and its functions. Whether in rooms full of closet-sized computers or in university halls, the dream of building intelligent machines has always been fueled by continuous research, studies and experiments. In the last ten years, AI has not only continued to be the object of study for scientists and academics, but has also gradually become part of the common, everyday public language: intelligent devices, Internet of Things and smart technologies are all recurring terms that, in one way or another, refer to certain capacities of technology to act intelligently.

Along similar lines, the pervasive diffusion of AI has influenced a wide spectrum of different domains: transportation, science, healthcare, education, communication and many more. This diffusion has gone hand in hand with the spectacle of the growing linguistic and logic potentials of AI. IBM's Watson on the quiz show *Jeopardy!* in 2011; Google's Deepmind AlphaGo in 2016; Elon Musk's OpenAI winning *Dota 2* tournaments in 2017, 2018 and 2019. In particular, the role of mass media has been central for the formation of a narrative line that has exalted the performative – but nevertheless opaque – features of AI. Accordingly, the growing enthusiasm for AI accomplishments has led to high expectations about its future advancements and its potential applications into society.

However, such expectations tend to overshadow other background mechanisms that are less spectacular and less exciting, and which involve many humans in their making. Despite the increasing capacity of AI technologies to automate more and more aspects of our lives seems to suggest that in the near future human functions will be less and less relevant for society, there is a large amount of invisible human labour involved underneath such developments, that tells a different story. Despite the rise of self-driving cars, autonomous delivery drones and robotics created a collective imaginary of AI as innovative and groundbreaking, there are several human practices behind the development of AI systems that contradict such views. These practices are rooted in the contemporary data-hungry paradigm of AI, for which the work of thousands of workers in labelling, curating, categorizing, correcting and sorting huge amounts of data is required: the classification of images for training computer vision in autonomous vehicles, the generation of audio files for training smart voice assistants, and many other practices that this thesis aims at exposing, allow the magic of AI to happen.

Accordingly, this thesis aims at providing a human-centered perspective on AI, to highlight how invisible forms of human labour shape the development of AI systems through practices of data curation and labelling. In order to achieve this goal, I will do an interdisciplinary review which includes literature on AI and machine-learning, Philosophy of Science and sociological studies on digital labour. Further, I will draw upon the literature on Critical Infrastructure Studies (Science and Technology Studies) to conceptualize the human infrastructure, thus providing a description of what it is, what the practices and forms of organization in which it manifests are, and how it finds a place within the context of AI. In order to map and trace the relations that the human infrastructure is entangled with, I will use the method of "infrastructural inversion", described by Bowker and Star (1999, 34) as a "struggle against the tendency of infrastructure to disappear". I will conceptualize infrastructural inversion as the human practices and arrangements which lie at the intersection with AI systems. Through this method, I will bring the human infrastructure out from the realm of invisibility, to firstly identify the reasons behind its own invisibility, and secondly, to intercept the ethical and epistemological issues that emerge from its becoming visible. In particular, I will discuss ethical concerns regarding workers labour conditions, and I will emphasize how a focus on the human infrastructure can provide us with conceptual and epistemological insights to evaluate how humans shape AI and its development.

This thesis acquires its importance in relation to the debate about the future working and ethical implications that AI systems will have on society. However, this research is not intended to situate itself in the debate, but rather to direct the attention of the debate towards pressing working and ethical issues which have not been voiced enough. Although the *development* of AI has led a lot of research to focus on important questions about the future of jobs, the potential impacts on society and the resulting ethical issues, fewer questions are asked about how *developing* AI is already reconfiguring the job market today, and how it is already leading to important societal and ethical issues[1]. While trying to anticipate the future, the risk is to lose sight of what is already happening in the present. In the *AI Now Report 2018*, an interdisciplinary team of researchers of the AI Now Institute highlighted some of the most pressing challenges due to the rise of AI technologies. Among the various strategies involved, the report mentioned two actions points, defined as "needed" for the future progress on AI-related issues: *infrastructural thinking* to better understand and track the complexities of AI systems, and *accounting for hidden labour* to call attention to the marginalized forms of human labour in AI systems (Whittaker et al., 2018). This thesis can therefore be situated between these two dimensions, to enrich the academic research concerned with these topics on one side, and to direct the public debate on AI towards issues that are little discussed on the other.

[1] Although the difference between "development of AI" and "developing AI" seems to be marginal, it has a clear scope for this thesis: while with "development of AI" I refer more generally to the global progress that has been made in the field of AI together with the application of AI technologies into society, "developing AI" emphasizes the gradual process by which such progress has been achieved. This distinction allows to highlight the complex system of interactions in which humans are situated within this process.

Thesis structure

To answer the research question “What is the picture of AI that emerges when it is analyzed through the lens of the human infrastructure that underlies its development?”, this thesis is structured in four chapters. In the first chapter, I will problematize the act of defining AI and the related issues that come with it. The fact that AI has not been defined until now invites a first, preliminary reflection on the wide variety of meanings and concepts that can be ascribed to this notion. It is possible to get a feeling for what is meant, even without providing a clear definition of what it actually is. I will therefore draw on the literature on AI and machine-learning, to analyze different ways in which AI can be defined. By doing so, I will firstly show how different definitions can frame AI in different ways, thus delineating what it is, and what it is not. Secondly, I will highlight how media can have an influence on the public perception of AI, by illustrating how it is generally defined in the public sphere, and how that circulates a certain image of what AI is. I will then provide a more adequate definition of AI which is relevant for the scope of this thesis, and which will allow to introduce the role of humans – and therefore that of the human infrastructure - in the context of AI.

In the second chapter, I will conceptualize and elaborate the notion of human infrastructure. First, I will introduce the field of Critical Infrastructure Studies (STS), from which I will draw theoretical and methodological tools to build my analysis of the human infrastructure. Scholars of the field have theorized about the various properties of infrastructures, but to sharpen the focus of my thesis I will mainly focus on some specific properties of infrastructures to conduct my analysis: *invisibility*, *embeddedness*, *reach*, *scope* and *scale*. I will base my methodological approach on the notion of “infrastructural inversion” as defined by Bowker and Star (1999). Infrastructural inversion means to recognize “the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other hand” (Bowker & Star, 1999, 34). This method operates as a “gestalt switch” (Bowker & Star, 1999, 34). It is a sudden change of perspective, which will allow me to bring to the foreground the network of arrangements, practices and organizations in which humans, as an infrastructure, are involved in the background. I will therefore describe the human practices of data labelling and curation and the forms of organization through which they are structured, to expose how these dimensions relate to the process of developing AI systems.

In the third chapter, I will deepen the infrastructural property of *invisibility* concerning the human infrastructure of AI, to show how the question of definitions and the focus on the human infrastructure as described in the previous chapters help to illustrate more clearly how workers invisibility occurs in the context of AI. I will therefore illustrate the dynamics that make and keep the human infrastructure invisible by focusing on the question of definitions and the role of digital platforms. By analyzing mechanisms of concealment that emerge in relation to specific practices and uses, I will expand and refine the concept of invisibility in relation to the human infrastructure.

After having exposed the various human practices, forms of organization and interrelations with technology, in the fourth chapter I will finally discuss the ethical, conceptual and epistemological issues that bringing the human infrastructure out from the realm of invisibility allows to address. I will discuss how the various forms in which the human infrastructure is organized affect the status of workers and I will highlight several ethical issues related to their labour conditions. Further, I will critically reflect on how a focus on the human infrastructure can provide a different conceptual perspective to look at the development of AI systems, in order to counter accounts which describe AI as highly automated and groundbreaking. Moreover, questions concerning bias and AI objectivity will be related to the human practices under scrutiny to discuss the epistemological issues involved.

Chapter 1
On defining Artificial Intelligence

Plato had defined Man as an animal, biped and featherless, and was applauded. Diogenes plucked a fowl and brought it into the lecture-room with the words, “Here is Plato's man!”. In consequence of which there was added to the definition, “having broad nails”.

-Diogenes Laertius, Lives of eminent philosophers.

Artificial Intelligence (AI) is famously hard to define. One of the reasons why this term is debated, is that the notion of *intelligence* is not easy to delineate in the first place. When thinking about humans for example, there are multiple forms in which intelligence can manifest itself: there are linguistic forms of intelligence, which involve the understanding of language and its different uses and nuances; there are spatial forms of intelligence, tightly related to the capacity of perceiving and interpreting the visual world; mathematical-logical forms of intelligence, involved in analytic and formal reasoning, such as the understanding of mathematical patterns, and many others (Gardner, 2011). This variety of forms suggests that a single, unifying notion of human intelligence is not only hard to come up with, but would also reduce the degree of complexity underlying intelligence. Similarly, it is unclear whether it would be possible – and nevertheless desirable – to provide a definition that captures the multiple nuances denoting intelligence in the field of AI.

Over the years, however, many attempts to define the concept of intelligence in relation to computers and machines have been made. Monett and Lewis (2018) have recently conducted a survey that contains more than 22 working definitions of AI, accompanied by other hundreds of suggested definitions from a cross sector of professionals and experts. Three well-known working definitions of AI are reported here:

- *“*Artificial Intelligence, the capability of computer systems to perform tasks that normally require human intelligence (e.g. perception, conversation, decision-making)*”* (David & Nielsen, 2016).

- “The art of creating machines that perform functions that require intelligence when performed by people” (Kurzweil, 1990).

- “Artificial Intelligence is […] the study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992).

Although these are only three among a multitude of other definitions of AI, one simple but significant observation about their content can be advanced: the first definition refers to AI as if it was a capability of certain computer systems; in this case, the capacity of perceiving, conversing or making decisions. The second refers to an art, according to which human-like capabilities can be reproduced from a machine. Lastly, the third relates to the study of computational systems capable of perceiving, reasoning or acting. The concept of AI is therefore framed in three different ways: AI as capabilities, AI as art and AI as a field of study. When this logic is extended to the larger amount of existing working definitions – which entails a larger number of conceptual framings – the notion of AI seems to become vague. In fact, a clear understanding of what is meant with the term AI fades in the vast amount of possible meanings surrounding this notion. Accordingly, if a single definition of AI is undesirable, and many definitions of AI confound the contours of its meaning, does this mean that AI cannot be defined?

Not necessarily. As Wang (2008) points out in *What Do You Mean by AI*, different definitions give AI different identities. In the field of AI, the working definitions of AI set the ultimate research goals to provide guidance and obtain valuable results. However, there is shared confusion among definitions of AI, to which different meanings are often implicitly ascribed from researchers. As a consequence, there are also different research goals, which require different methods, and which produce different results – evaluated through different criteria – thus resulting in a fragmentation within the field. In fact, definitions are not all the same, and they often hold different underlying presuppositions. Each definition can indeed illuminate some aspects of AI while obscuring some others, portraying a specific picture of what AI is. To this regard, Wang points out that when it comes to evaluate the similarities between the intelligence of humans and computers, 5 typical ways to define AI can be distinguished. I summarized them in the following table:

Definition	Description	Examples
Structure-AI	It requires the structural similarity between an AI system and the human brain. AI can be achieved by building a brain-like structure, consisting of massive neuron-like processing units working in parallel.	Artificial Neural Networks

Behavior-AI	It requires the behavioral similarity between AI system and human mind. AI is evaluated by testing systems' behavior.	Turing Test
Capability-AI	It requires an AI system to have human capability of practical problem solving. The intelligence of an AI system is indicated by its capability of solving hard problems.	Chess-playing system DeepBlue
Function-AI	It requires an AI system to have cognitive functions similar to those observed in humans. AI is represented as a function that maps input (percepts) into output (actions).	IBM's Watson
Principle-AI	It requires an AI system to follow similar normative principles as the human mind. It aims at identifying the fundamental principle by which human intelligence can be explained and reproduced in computers at a general level.	None

Fig.1 - Typical ways to define AI. From Wang (2008) *What Do You Mean by AI?*

Although these types of working definitions all set legitimate research goals (Wang, 2008, 7), they carry a range of assumptions that cannot be ignored: for example, the definition of AI by Principle implies that there is a fundamental law by which human intelligence can be explained; once this law would be discovered, it would be possible to reproduce it into computers. A Capability-AI definition, on the other hand, identifies an agent as intelligent in relation to its capability to solve hard problems; whether it shares fundamental human principles or not, is irrelevant for its definition of intelligence. This means that when it comes to evaluate whether a system such as DeepBlue – the rule-based[2] computer programmed to play chess – is intelligent or not, the matter is one of definition. According to Capability-AI types of definition, the system would be classified as intelligent since it is capable of human-like problem solving abilities. Contrarily, according to Principle-AI definitions, DeepBlue would not be defined as intelligent, since the way it is programmed to function does not replicate the (unknown) principle underlying human intelligence. The answer is therefore derived from the implicit assumptions that each definition carries with it. This brief example shows that the question of defining AI is not merely a matter of choosing the best definition among many others. What it rather indicates, is that each definition plays a substantial role in delineating what stays *in* and what stays *out*; according to different definitions, technologies can, for instance, be classified as intelligent or not.

[2] Rule-based programming consists in a set of rules that tells the system what to do or what to conclude in different situations, miming the reasoning of human actors (Grosan & Abraham, 2011, 149).

1.1 Portrayals of AI in the public sphere

Different definitions can frame AI in different ways. Research has shown the relation between people's beliefs and impressions about AI and the media's views on it (Cave et al., 2018; Chuan et al., 2019; Fast et al., 2017). The way media define AI can therefore influence how it is publicly perceived. While Wang (2008) problematizes the role of definitions in the field of AI, arguing that the shared confusion about its meaning has negative effects on the research outcomes, in this section I will point out that this confusion affects also the public sphere. In particular, I will illustrate the two paradigmatic ways in which media generally define AI, as a technological application and as an entity, to explain why they are inadequate to understand the central role of humans in the context of AI.

A research conducted in the UK by the Reuters Institute (2018) reveals that nearly 60 percent of news articles reporting on AI are indexed to industry products, announcements or initiatives (p.1). Almost two thirds of the articles referring to AI are framed around industry products, which the report claims ranging from smartphones and running shoes, to sex robots and brain preservation (p.3). As a consequence, media outlets generally define AI in terms of specific technological applications, which somehow *are* or *possess* AI: self-driving cars, voice assistants, smart wearables. This view on AI *as* a technological application defines it solely in terms of specific technological artefacts, and is clearly exemplified in news articles headlines like "Data from wearables helped teach an AI to spot signs of diabetes" (Engadget, 2018), or "Google's Artificial Intelligence Built an AI That Outperforms Any Made by Humans" (Futurism, 2017), in which the article "an" before the noun already frames AI as if it was an actual thing.

Portrayals of AI are also part of fictional (popular science fictions, imaginative thinking about future intelligent machines) and non-fictional (media coverage about AI and its effect) narratives, which can be disconnected from the reality of the technology, since they either focus on scenarios that are decades away from becoming actual, or are just part of a small subset of issues within the larger field of AI (Cave et al., 2018, 14). A Royal Society (2018) report says in this regard that "Popular portrayals of AI in the English-speaking West tend to be either exaggeratedly optimistic about what the technology might achieve, or melodramatically pessimistic" (p.9). High expectations and false fears about AI and its effects on society can be attributed to these kinds of narratives that, contrarily to those which define it as a specific technological application, often only vaguely define AI, or do not define it at all. As a consequence, the concept of AI remains very abstract, sharing the characteristics of an *entity*, of which contours are unclear and of which capabilities are generally over-estimated.

Both ways of defining AI, as a technological application and as an entity, have an effect on the public sphere, which is characterized by the mundanity of everyday, large-scale information that

I have claimed being particularly relevant for framing people's perceptions and beliefs about AI. In fact, these views fail to provide people with an adequate representation of AI. Moreover, the quality of the information provided is compounded by the fact that media's coverage often lacks the opinions and engagement of experts and informed decision makers (Dubljević, 2012; Grant et al., 2011). As a technological application, the concept of AI is limited to - and framed only in relation with - specific automated technologies, which are attributed as artificially intelligent agents. As an entity, the public discussion is distorted by the polarized "hype and hope" and "gloom and doom" perspectives on AI (Dubljević, 2012) which are neither informative nor explanatory, but form over-optimistic or over-pessimistic views on AI.

Either way, both definitions as a technological application and as an entity presuppose the existence of an independent capacity of AI to intelligently act by itself. Both ways of framing it, share indeed the assumption that AI possesses automated capabilities, whether in the form of technological artefact or entity. Automation is indeed one of the primary qualities of AI, which makes it such a powerful driving force of change and disruption. However, automation has both a technological dimension and an ideological function (Taylor, 2018). The technological dimension is represented by the actual capacity of AI-based technologies to independently act and perform tasks in the world. The ideological function, on the other hand, represents the set of narratives, beliefs and values usually attributed to AI. This ideological function tends to "oversell" (Taylor, 2018) automation, in the sense that the capacities of automated technology are typically exaggerated, thus reflecting a distorted picture of AI. In particular, the widespread belief that humans will be less and less relevant in various aspects of society due to the rise of automated technology, is consistent with the ideological function of automation. This view is epitomized in the notion of "useless class", which Harari (2017) defines as a mass of economically and socially irrelevant people that will not only be unemployed, but will be unemployable due to their lack of competences in the face of the rise of algorithms and AI technologies. These views have far-reaching consequences on how AI is publicly perceived and therefore on how people evaluate its capacities. But more importantly, these accounts do not give any relevance to the role of humans in the context of AI. On the contrary, they completely exclude them from it. To counter views of AI as a technological application or as an entity, which reinforce inadequate representations of AI, I will provide a more realistic perspective on it, to challenge accounts which attribute unrealistic automated and disruptive properties to artificially intelligent agents. In doing so, I will emphasize the utterly central role of humans in this context.

1.2 Providing a definition of AI

“It has been suggested by some that as soon as AI researchers figure out how to do something, that capability ceases to be regarded as intelligent - chess was considered the epitome of intelligence until Deep Blue won the world championship from Kasparov - but even these researchers agree that something important is missing from modern AIs” (Bostrom et al., 2014, 3). In this passage from *The Ethics of Artificial Intelligence*, Bostrom and Yudkowsky suggest that what is considered to be intelligent in the field of AI, changes according to what machines are capable of accomplishing. This implies that the more machines learn how to perform new tasks, the less intelligent the previous, old tasks seem to be. The variable notion of intelligence is indeed tightly dependent on historical and evolutionary circumstances: what was called AI yesterday, may no longer be today. Consequently, to provide a good definition of AI and unveil how humans are involved in this context, it is firstly necessary to delineate what is regarded as (artificially) intelligent today.

The two paradigms that mark the clearest distinction between the past and the present in the field of AI can be presented as Symbolic AI or GOFAI (Good Old Fashion Artificial Intelligence) and Connectionism[3]. The first, was the predominant approach until the late 1980s; it is called symbolic since programming involves the manipulation of symbols, intended to represent concepts that refer to objects in the external world (Willshaw, 1994, 87). It is rule-based, which implies that the series of logic-like reasoning steps that symbolic AI systems carry out, follow from a formally specified set of rules encoded into the computer program (Garnelo et al., 2019, 17). The capacity of a machine to learn is therefore limited to the set of rules programmed in it.

Connectionism, on the other hand, is the predominant approach to AI today; it is inspired by the anatomy and physiology of the nervous system, of which models usually take the form of neural networks (Barrow, 1996, 135). This approach does not involve the direct manipulation of symbols, but the capacity to learn lies in the connections of the networked structure of the models (Bereiter, 1991, 12). In fact, contrarily to Symbolic AI, there is no specific set of rules to be rigidly followed: machines are programmed to learn from past experience and data (Alpaydin, 2010, 3). Or to be more precise, from huge amounts of data. Over the last decade, this paradigm has progressed along with the exponential growth in data production and computing capacity for storing and processing large amounts of data. The most recent technological breakthroughs in the field of healthcare, transportation, communication or science have been possible because of data availability, rather than encoded rules. Intelligence today, is data-driven. But more importantly, it is determined by the methods through which data are manipulated. AI can thus be defined as “a set of computer science

[3] Sometimes called non-symbolic AI or sub-symbolic AI.

techniques that enable systems to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making and language translation" (McCauley, 2016, 3). Framing AI as the actual methods employed today to manipulate data, allows to unveil the more implicit human practices involved, which will provide us with a different perspective to look at AI and its development.

1.3 Introducing the role of humans in developing AI

As stated, the progress and evolution in the field of AI depends on the set of computer science techniques and methods that guide it today. The research has enormously improved thanks to the progress made in a branch of the field called machine-learning, which aims at teaching machines how to learn from data. Here, I analyze the basic structure of machine-learning, to show how and where humans are involved in data processing for developing AI. This will allow me to set the context in which to introduce the concept of human infrastructure.

The basic structure of machine-learning consists in finding a mathematical function (g*)* that correctly maps the relationship between a set of inputs X (x_1, x_2..) and its corresponding set of outputs Y (y_1, y_2..). The function (g) is an approximation of the target function (f), which is unknown and represents the correct mapping relationship between the set of inputs X and its corresponding outputs Y (Abu-Mostafa et al., 2012; Karaca, 2019). In order to find the function (g), a large set of data (also called training data) is needed. The training aims at finding regularities, patterns and structure in those data in order to build a mathematical model (Nasteski, 2017, 53). A key element for the development and use of machine-learning models is indeed the elaboration of large amounts of data: without enough data, it would be impossible for machine-learning models to be trained properly, and therefore for autonomous cars to drive, or for voice assistants to answer questions about the weather. After being trained, the model would then be applied to new data sets and evaluated according to its capacity to make correct predictions in different applications.

For example, if a company would automate the candidates hiring process with machine-learning techniques, the process would be as follows: the function (g) to find would approximate the target function (f), namely the true representation of the relationship between the set of inputs X (for example the age, working experience, educational level of the candidates) and the set of outputs Y (for example being classified as a potential candidate to hire or not). In this case, the aim would be to construct a model that will be used to automate the hiring process. The model would be trained with a dataset of thousands of sample data containing age, working experience and educational levels of various candidates. After the training, the model would be applied in practice to classify new, potential candidates on the base of their characteristics, and finally, a right prediction would mean to do that correctly.

The method that I have just described represents the most popular type of learning, which belongs to the supervised machine-learning paradigm[4]. Supervised learning is "the most studied and most utilized type of learning" (Abu-Mostafa et al., 2012, 11). It is indeed one of the dominant methodologies in machine-learning (Nasteski, 2017, 60), and implies the use of large sets of *labeled* data. However, what is often bypassed in the discourse on AI is the unexamined nature of the term *labeled* before *data*. Labeled data can be defined as data to which one or more pieces of information, or labels, are attached. Labels are key features (such as characteristics or properties), attributed to unlabeled datasets, that are needed for machine-learning models to identify patterns and structures among data; labels provide data with a target, which determines what is the kind of output (or correct answer) that the machine-learning model will have to predict (CloudFactory, 2019, 4). That of data labelling is an act of classification, that far from being automated, requires human functions. If we were to situate the role of humans in the aforementioned machine-learning process scheme, it would be placed in it as follows:

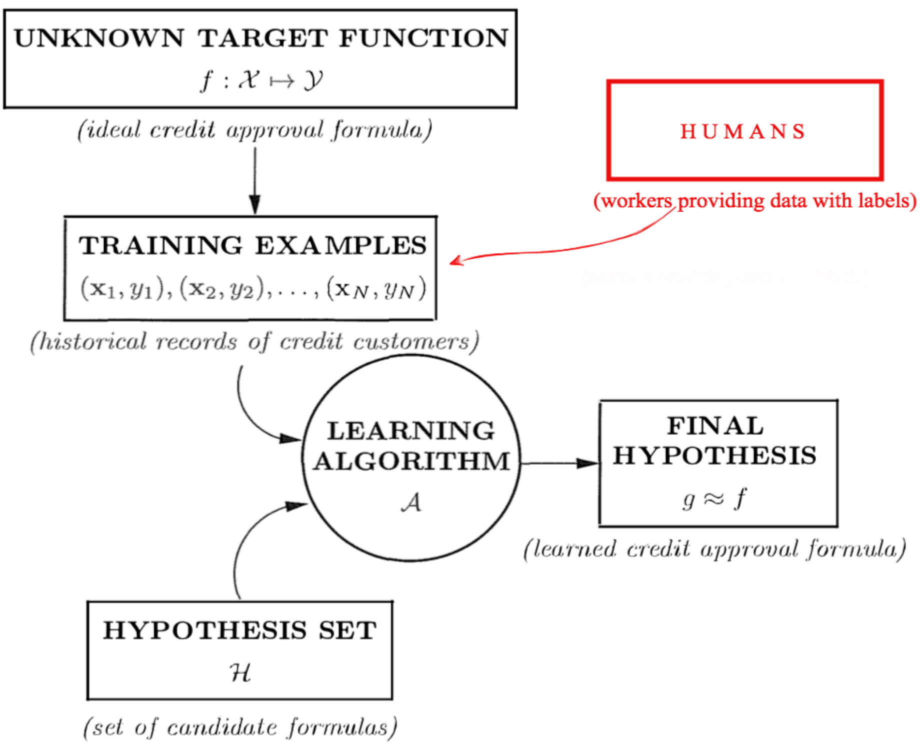

Fig.2 - Slight modification (in red) of the basic learning setup scheme provided by Abu-Mostafa et al. (2012).

[4] I acknowledge that there are two other paradigms of machine-learning (unsupervised learning and reinforced learning). However, their use today is not as widespread as that of the supervised learning paradigm. The most relevant difference between supervised and unsupervised learning, is that supervised learning datasets contain explicit examples of what the right output should be for the given inputs; for unsupervised learning, the dataset does not contain any output information (Abu-Mostafa et al., 2012), data are therefore called *unlabeled* or *raw*.

Although this labelling function of humans is neither popular nor recognized as fundamental in the AI debate, it plays a key role in enabling machine-learning models to learn from data. Labelling data is a time-consuming task which relies on the manual and cognitive abilities of humans to be performed. However, the ideological function of automation tends to glorify a narrative line that highlights the independent intelligence and automation capabilities of AI systems - thus obscuring, excluding and limiting the relevance and purpose of humans.

I have until now highlighted the issues related to the act of defining AI, to show that definitions carry different underlying assumptions, which can illuminate some aspects of AI while obscuring some others. Then, by illustrating how AI is generally defined by media, as a technological application and as an entity, I explained that these framings affect the formation of people's beliefs and opinions about AI in a way that does not allow to further explore the role of humans behind the ideological function of automation. I then provided a definition of AI which is adequate for the scope of the thesis, and I analyzed the basic structure of machine-learning to point out where human functions are situated in it. It is now from the role covered by humans that the next chapter unfolds. By introducing and elaborating on the concept of human infrastructure, I will explain how humans not only shape the development of AI, but also constitute the fundamental infrastructure upon and through which its progress is made possible.

Chapter 2

The Human Infrastructure of Artificial Intelligence

This chapter introduces and elaborates the concept of human infrastructure. Developing this concept is useful to situate humans in the context of AI and to show how central is their involvement in the process of data labelling and curation. The goal is to provide a conceptualization of humans *as* an infrastructure, to provide a human-centered perspective on AI that allows to raise and discuss critical ethical concerns in regard to labour conditions and AI-related epistemological issues. I will therefore describe the practices of data labelling and curation with which humans as an infrastructure are involved, and the forms of organization through which these practices are structured. In order to do that, I will firstly describe the main field of Critical Infrastructure Studies, on which I will build my theoretical and methodological analysis of the human infrastructure.

2.1 Critical Infrastructure Studies

Critical Infrastructure Studies is a field of study that aims at investigating infrastructures, their evolution over time, and the multiple ramifications in which they unfold in space. Scholars of the field have worked to identify and clearly delineate the properties of infrastructures. Here, I will point out the main properties of infrastructures that I will consider to characterize the notion of "human infrastructure" and to describe the ways in which its interactions take place in the context of AI. To sharpen the focus of the thesis, among the many existing properties of infrastructures, I will mainly focus on *invisibility, embeddedness, reach, scope* and *scale*, which are the most relevant for my analysis. In this analysis, I will integrate these properties together with the core method of "infrastructural inversion", which will be used to trace relations and to shift to the foreground the human infrastructure that invisibly operates in the background.

The study of infrastructures has been cultivated within different fields of study, such as history, anthropology, social sciences and Science and Technology Studies (STS). The term "infrastructures" has been widely spread since the '90s in various areas through journalism, governments, information systems and academia (Edwards et al., 2009, 365). Infrastructures are usually referred to as a system of substrates - like railways, electrical power plant, wires and cables, pipelines, plumbing etc. (Star, 1999, 380), or more broadly as "material forms that allow for the possibility of exchange over space" (Larkin, 2013, 327). Both ways of addressing it, encapsulate the tendency of thinking about infrastructures as something exclusively material, which is how they are generally addressed. However, along with the crucial relevance of their material and physical features, infrastructures have been framed in the field of Infrastructure Studies mostly in reference to the methodological and conceptual tools they offer. In fact, studying infrastructures does not simply involve the study of physical networks. Instead, it is from the branched structure of physical networks themselves (just think of the connective structure of electricity grids and highways), that conceptual material is offered to the study of infrastructures; the notion of infrastructure evokes images of interconnectedness and interdependence which escape the rigid limits of physical structures. Rather than infrastructures themselves, it is the various ideas, concepts and forms of abstraction that originates from the physicality of infrastructures that matter for inspiring and creating new ways of theorizing about networks.

2.2 Properties of Infrastructures

The intrinsic *relational* property of infrastructures makes them a very powerful exploring tool: since they cannot exist in isolation, but are "inextricably linked to other technological, social, political, and economic actors, networks, and processes" (Ensmenger, 2018, 14), infrastructures - other than physical networks - represent the perfect methodological tool to identify relationships, links and connections that constitute the phenomenon under scrutiny. So far, a few properties of infrastructures have been mentioned, without considering the word itself: "infrastructures" literally means 'those structures that are below' (Pasveer et al., 2018, 6). It refers to something below a surface, like the plumbing pipes in the wall. But, in a more abstract sense, it also refers to what is below our perceptions and investigations (Pasveer et al., 2018, 6). "A good infrastructure is hard to find", claim Bowker and Star (1999, 33). Infrastructures are ‘‘by definition invisible, part of the background for other kinds of work’’ (Star, 1999, 380). There are hundreds of cases that exemplify infrastructures *invisibility*: just think about every day, simple actions such as filling a glass of water or turning on the light. A vast network of plumbing, wiring and distributions grids is in action, although invisible to our eyes. Despite its usage, it is not directly to the infrastructure itself that one's attention is directed, but rather to the task that the infrastructure allows to perform. Invisibility could be therefore intended as being a property of infrastructures.

However, various authors have been deeply engaged with the concept of infrastructures, proposing more nuanced views on what they are, and how their invisibility can be better conceptualized. Infrastructures are not just something hiding in the background and ready to be used. They rather represent the space in which multiple practices, work, people, things, information and routines unfold and converge. That of infrastructures can be thought as a space of flows (Castells, 1996). Accordingly, in order to understand how AI, humans and labels come together within this concept, more layers of what infrastructure are, need to be deepen. As a starting point, rather than thinking about *what* infrastructures are, Star and Ruhdeler (1996) propose to think about *when* infrastructures are: since infrastructures are never just something isolated, but always relate to various activities in different contexts, geographies and structures, they are better conceptualized as something that emerge in practice. “Infrastructure is a fundamentally relational concept, becoming real infrastructure in relation to organized practices” (Star, 1999, 380). Put in these terms, the image of infrastructures seems to acquire motion; asking *when* rather than *what*, emphasizes the dynamic character of infrastructures, that far from being just a motionless network in the background, emerge as the junction where an orchestra of multiple relations unfolds. Accordingly, also the related notion of invisibility previously identified as a given property of infrastructures now becomes situated, varying in accordance with the contexts. In fact, infrastructures are not invisible for everyone, but have different degrees of visibility according to different people and situations. As Star (1999, 380) convincingly put it, “for a railroad engineer, the rails are not infrastructure but topic” and “the cook considers the water system as working infrastructure integral to making dinner. For the city planner or the plumber, it is a variable in a complex planning process or a target for repair”.

As defined by Star and Rudheler (1996), infrastructures embody also other dimensions which play a crucial role for the subject under investigation. One is *embeddedness*: being embedded means that infrastructures are often sunk into other organizations, technologies and social configurations. To be inside other structures, can be thought of as a consequence of another dimension of infrastructures, according to which they are *built on an installed based*. This means that every infrastructure, instead that out of nowhere, is always built on another base. For example, the fire alarm infrastructure of a building can be thought of as part of the electrical infrastructure, which in turn is part of a larger infrastructure composed of walls, floors, foundations and so on. Naturally, each of these infrastructures is always entangled with other social (but also political, legal and economic) ones, composed of a thick network of policies, safety regulations, standards, rules and so on. Nevertheless, as Edwards et al. (1996) nicely put it, it is inaccurate to think about infrastructures as something that is being *built*; using instead the metaphor of *growing* an infrastructure, they capture “the sense of an organic unfolding within an existing (and changing) environment” (p.369). In this sense, single infrastructures emerge as part of a whole by leaning on other existing ones.

Another crucial property of infrastructures for this analysis is *reach* and *scope* (Star et al., 1996, 113). The main idea behind this dimension is that infrastructures can extend beyond their on-site presence. Reach and scope are two variables setting the boundaries and contents of infrastructures: reach can be thought as the amount of processes and activities that are touched by an infrastructure, while scope as the variety and type of applications that can run on it (Ciborra et al., 1998, 307). The notion of reach and scope is strongly related to that of *scale*; the literature on infrastructures calls attention to multiple varieties of scale, such as that of time, force, size, space or social organization (Edwards et al., 1996; Edwards, 2003). However, scaling infrastructures usually refers to making systems bigger and extending their reach (Edwards et al., 1996, 370). This process of extension always implies the relation between two dimensions: local and global, which relationship can be conceptualized in two ways[5]. The first, as scaling-up: a movement extending from the local (particular, small, individual) towards the global (general, large, collective). The second, as local/global being an interpretative framework to analyze how situated, local practices and activities relate to larger tendencies and dynamics within infrastructural dimensions.

2.3 Infrastructural Inversion

In order to proceed with my analysis, I will draw on the conceptual method of *infrastructural inversion*. That of infrastructural inversion is a method defined by Bowker and Star (1999) as a "struggle against the tendency of infrastructure to disappear (except when breaking down). […] Infrastructural inversion means recognizing the depths of interdependence of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other hand" (p.34). This method entails to carefully observe the processes that are often considered to be boring, behind the scenes, in the background, and bring them to the foreground (Bowker et al., 1998, 234). The method of infrastructural inversion is used here with a specific focus on the human infrastructure, to bring to the foreground the arrangements and activities involving humans in the process of data labelling and curation, thus uncovering the interdependencies between the development of AI for which these practices are needed and human labour. The contraposition between background/foreground, invisible/visible, implicit/explicit is central for understanding the dimension in which the method of infrastructural inversion operates. More importantly, this conceptual method allows to expose the human infrastructure and the depth of its interconnected relationships. As Edward put it: "To understand an infrastructure, you have to invert it. You turn it upside down and look at the 'bottom' – the parts you don't normally think about precisely because they have become standard, routine, transparent, invisible." (Edwards, 2010, 20). Having illustrated the most relevant properties of infrastructures for the scope of this thesis and the methodology of infrastructural inversion, I will now introduce the notion of "human infrastructure", to which all these infrastructural dimensions will be integrated.

[5] For a more detailed account of the relation between local and global, see "Gibson-Graham, J.K. (2002) 'Beyond global vs. local: economic politics outside the binary frame', in A. Herod and M.W. Wright (eds) Geographies of Power: Placing Scale. Oxford: Blackwell, pp. 25–60".

2.4 The Human Infrastructure of Artificial Intelligence

All the dimensions mentioned above come together as useful methodological and conceptual tools to analyze, explore and map the human infrastructure that underlies the development of AI. However, it is necessary to first develop the concept of human infrastructure and explain how it is configured within the context of AI. The concept of human infrastructure incorporates two apparently separated notions, that of *human* and that of *infrastructure*. Having already pointed out some features of infrastructures, I will now explain how humans can be defined *as* an infrastructure in itself - hence, human infrastructure - on a conceptual level. After that, I will turn my focus to the activities and tasks involved within it, to show why humans are an infrastructure *in practice*.

Edwards (2003) suggests that one way to think about infrastructures is by doing that negatively, namely as “those systems without which contemporary societies cannot function” (p.187). This formulation can be used to highlight the first reason why humans can be conceptualized *as* an infrastructure, that is the constitutive, fundamental element without which a system - in this case the one through which AI progress has been made possible – could not exist. Framing humans as infrastructure, aims to bring back to humans the attention and relevance that is often, in one-direction, channeled towards technological advancements and applications in the field of AI. In particular, it is a reminder that without humans, those achievements would not be possible. Trivial as it may seem, the act of highlighting the role of humans in developing AI is a fundamental point that risks being easily overlooked.

The second reason to conceptualize humans as an infrastructure, is that infrastructures are commonly associated with physical structures like electric grids and railways, and not with people (Mateescu et al., 2019, 13). By framing humans as infrastructure, humans are metaphorically reduced to objects; this expression, encapsulates “the tension between calling out humans as infrastructure and the reduction of human to infrastructural object [6]” (Mateescu et al., 2019, 13). This metaphor reflects the limits of human expression that working *as* an infrastructure entails, and represents the reduction of humans to mere inanimate parts of a larger system.

Third, and most useful, this conceptualization allows to ascribe the properties of infrastructures to humans, opening multiple ways to analyze and discover their position within the context of AI. Some of the characteristics of infrastructures have already been mentioned, but one of them is particularly relevant here: infrastructures tend to fade into the background, becoming

[6] I first (and only) encountered the notion of “Human Infrastructure” in Mateescu, A., & Elish, M. C. (2019). AI in context: The labor of integrating new technologies. Data & Society report. Despite they focused their research in the context of AI and Farm Management & Grocery Retail, I retain that this concept would benefit from a further elaboration and application in more areas of research.

invisible. Invisibility is a necessary condition for an infrastructure to work well, but is usually interrupted when the infrastructure breaks down (Star et al., 1996, 113): for example when there is a blackout, the Wi-Fi stops working, or a pipe starts spilling water. While in this occasion infrastructures become visible, their natural tendency is to disappear. As an infrastructure, the human labor that underlies the development of AI tends to fade into the background. It becomes invisible. Conceptualizing humans as an infrastructure allows to firstly recognize that they are not visible, and secondly allows to explain how and why their invisibility occurs. Up to this point, the analysis of the human infrastructure has been mainly addressed from a conceptual and methodological point of view. From the next section onwards, by focusing in detail on the multiple tasks, practices and activities implied in developing AI, I will explore and map the role of the human infrastructure in practice.

2.5 The Human Infrastructure in practice

Until now, the human infrastructure has been mainly addressed from a conceptual and methodological perspective. From now on, I will integrate this concept into more concrete and practical dimensions by diving into empirical research. The methods pointed out so far will be therefore used to guide the empirical analysis of the human infrastructure, to provide an account that considers the multifaceted ways in which the human infrastructure manifests itself. The goal is to explore and understand what the human infrastructure of AI is in practice, to explain how humans shape AI with real-world examples and cases. This means looking at the actual tasks, activities and processes involved in developing AI, to understand how and according to which dynamics are humans situated in it.

A first helpful, preliminary step to identify the configuration of an infrastructure, is to look at tensions. Looking at tensions is a common practice in the field of Infrastructure Studies to reveal the "conflicting goals, purposes and motivations" (Ribes et al., 2009, 376) of actors and participants involved in the development of infrastructures. Looking for tensions facilitates the identification of an infrastructure, thus making it visible to see what it entails, and for whom. Tensions are particularly useful to observe in the moment of formation of infrastructures, during which intense conflicts are involved; in these moments, "the identity and status of relevant stakeholders, the distribution of benefits and losses, and the general rules of the game are all being worked out simultaneously" (Jackson et al., 2007).

Accordingly, to start seeing how the human infrastructure manifests itself in the context of AI, a way of doing it is by looking at tensions: in accordance with the idea that infrastructures have inherent relational properties, looking at tensions unravels the *when* of an infrastructure, namely the precise circumstances under which the infrastructure began to grow. The most relevant tension to pay attention to in this case, is one of *scale*. The tension of scale I am referring to, is located precisely between two moments: the exponential growth of data production over the last few decades, and the need to have those data labeled. As mentioned before, in order to train supervised machine-learning models - the hearth of AI growth - data need to be labeled, and that still requires relevant human functions. Labeled data are often defined as the "bottleneck" to the growth of AI industry (Ratner et al., 2020; Chew et al., 2019; Roh et al., 2019;), because their scarcity slows down and hinders the whole process of technological innovation. A lot of data circulates within this paradigm, but data without labels is almost useless.

Therefore, there is a significant gap between the huge amount of data produced, and the scarce number of labels attributed to this data; and it exactly inside this space that the human infrastructure has been invisibly growing. The identification of this tension allows to see more concretely where the human infrastructure is situated within the context of AI, and to highlight the functions that it covers. At the same time, it hints at the extension and relevance of the infrastructure, suggesting that humans do not just “fill the gap” in a system, but they rather represent the fundamental component without which that system could not work. Taken together, these reasons slowly begin to delineate the first contours of the human infrastructure of AI in practice. To mark these contours more clearly, I will in the next section dive more deeply into the functions of the human infrastructure by unpacking the notion of “labelling”. The main objective is to illustrate the tasks, activities and processes in which humans are involved in developing AI to substantiate their relevance.

2.6 The practices of the Human Infrastructure

In the previous paragraphs it has been sketchily defined what labeled data are, and why they are fundamentally needed to transform raw data into usable ones within the machine-learning paradigm that leads AI developments today. In this section I will describe more in detail the practice of labelling, to show what it is and how it takes place in the context of AI. Data labelling belongs to a larger set of practices of data curation and data manipulation, which involve activities of content moderation, images segmentation, audio and document transcription that will be addressed in this section. Unpacking the notion of labelling helps to explain how people relate to the act of labelling, and how they are accordingly configured within the larger infrastructure they constitute in practice.

Labelling[7] is an act of classification involving human cognitive functions; it is a concept that sums up in itself a heterogeneous variety of tasks and functions. It is sometimes more roughly referred to as part of the “Human-In-The-Loop” (HITL) model, a feedback system used in machine-learning to indicate the role of humans inside the chain of processes that lead to a final model or application.

[7] Labelling can be interchanged with “annotation”.

Accordingly, unpacking the notion of labelling also means to specify the general notion of Human-In-The-Loop, to understand how humans are involved in the loop, and in what the loop precisely consists of. In *The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence*, Tubaro et al. (2020) distinguish three poles in which human functions occur within the larger process of machine-learning development: AI *preparation*, *impersonation* and *verification*. I will use these three different categories to illustrate and organize the various tasks and functions covered by the human infrastructure to explain how human knowledge is transferred into machines. In this way I will show, using the global/local interpretative framework of scale, how the global development of AI - and the consequent human-related capacities of machines to interpret, structure, match, diagnose, discover, etc. (Boon, 2020) - is made possible through local practices of labelling.

AI Preparation

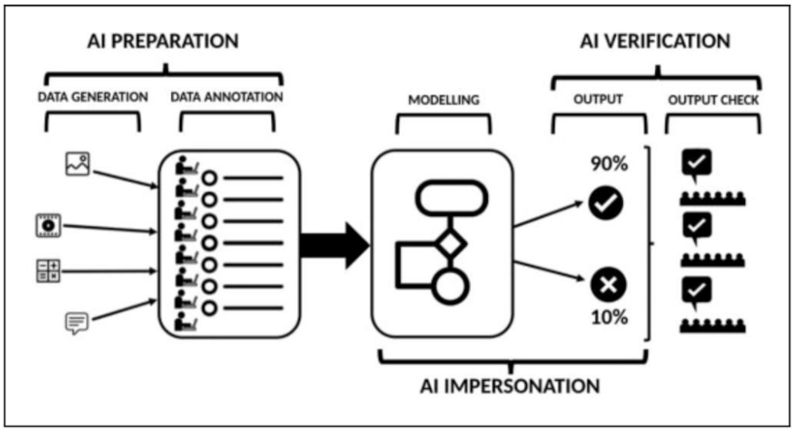

Fig.3 - Three human functions in the development of machine-learning based AI solutions as described by Tubaro et al. (2020).

AI *preparation* represents the primary phase of the paradigm, divided by Tubaro et al. (2020) in two parts: data *generation* and *annotation* (or labelling). Data generation, as the term suggests, involves humans in the generation of data, that are then collected for training machine-learning models. One of the most common examples of data generation is audio utterance collections. In this case, the generation of audio data comes in the form of voice recordings: data are gathered by having a large amount of people recording and repeating a few short predefined sentences, in which a variety of vocal timbres, accents and uses of slang are collected (Tubaro et al., 2020, 5). In particular, the relevance of the *local* task of recording one's own voice, emerges in relation to the *global* development of AI-based technologies. Worldwide spread AI technologies like Apple's Siri, Microsoft's Cortana or Amazon's Alexa, are indeed directly concerned with the content of these practices. In fact, the heterogeneous variety of recorded voices and timbres is mainly employed to design smart voice assistants. To have an idea of the scope, the use of these devices is expected to increase to 4.2 billion units by the end of 2020 (Juniper Research, 2020). Even if Tubaro et al. distinguish here between data generation and labelling, I sustain that also data generation is a form of labelling: labeled data are data to which one or more pieces of information (or labels) are attached. In the case of voice assistants, the sentence to be read is a form of raw data, emptied from any kind of significance. The vocal audio, in turn, represents the additional information that is attributed to the sentence, which becomes in this way labeled - and therefore usable for machine-learning models to identify patterns and structures among data. One can for example read from Alexa's FAQ: "Alexa is designed to get smarter every day. [...] This training relies in part on supervised machine-learning, an industry-standard practice where humans review an extremely small sample of requests to help Alexa understand the correct interpretation of a request and provide the appropriate response in the future. For example, a human reviewing a customer's request for the weather in Austin can identify that Alexa misinterpreted it as a request for the weather in Boston" (Amazon, 2020). In this case, users' voices recorded by Alexa are analyzed by humans to control whether the machine has correctly understood the sentence - and eventually fix it. With data generation, the process is similarly inverted: the correct sentence is what humans are provided with, and their task is to pronounce it in a way that matches the written text - thus sticking a label. In one case or another, in order to have functioning voice assistants, a lot of humans are involved in the loop, but their role is hardly visible: what is instead brightly apparent, is the seemingly magical ability of voice assistants to converse and answer questions about the weather.

Data annotation (or labelling) represents the second part of the preparation pole and encompasses a huge variety of specific tasks and practices, that find application in a lot of different spheres within the global development of AI. It represents a core practice in almost any context in which supervised machine-learning methods are applied and consists in the classification of huge varieties of audio, video, image and text data. It is an essential step in the shaping of AI *sight* in the area of computer vision, *hearing* in the area of speech recognition, and *language understanding* in that of natural language processing. The most common applications in which this practice is involved, are the automotive industry (e.g. self-driving cars), aerial imagery (e.g. drones vision), augmented and virtual reality (e.g. object and sentiment recognition), but also retail and e-commerce (e.g. autonomous check-out, theft prevention), robotics and manufacturing (e.g. logistic management, inventory handling), and many others (ScaleAI, 2020). If we take the area of computer vision, labelling concerns primarily the classification of images and videos. In the field of self-driving cars for example, the main objective of applying computer vision based on machine-learning is that of teaching autonomous cars to see - or better, to see properly in order to avoid fatalities. Surprisingly, however, it is not inside of software, computer programs or intelligent algorithms that the raw source of this knowledge can be found. Contrarily, it lies in the cognitive functions of a myriad of human labelers. While the debate on self-driving cars is mainly focused on whether an autonomous car should invest X or Y (the classic ethical trolley problem), less discussion is focused on how the vehicle distinguishes between X and Y in the first place. The process of image recognition, according to which an autonomous car can discern between a road lane and a sidewalk, involves indeed a painstaking work of hand-made labelling. This handiwork concerns almost anything that a car may encounter on its way: trucks, pedestrians, cyclists, traffic lights, road signs, road lanes, cats, trees, strollers, trash cans, and any other relevant object contained in huge databases of images and videos. Each of these objects needs to be classified, which means that they must be carefully sorted out and outlined. There are various methods to do that, which may vary in relation to the goal to achieve. In the following figure, some of them are reported:

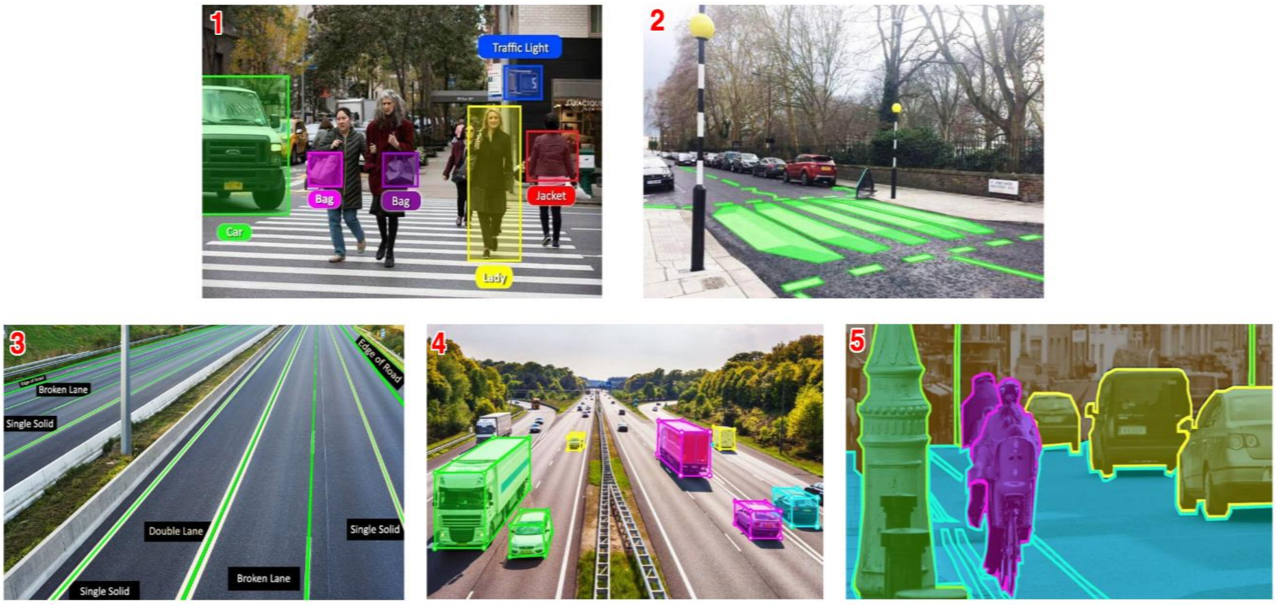

Fig.4 - Some examples of labelling methods for self-driving cars (Anolytics, 2020).

Each of the methods illustrated above is employed for specific functions:

- The method used in image n.1 is called "2D bounding box" and is designed for object detection (e.g. pedestrians, traffic lights, etc.); here, the person doing the labelling has to draw a box around different objects and specify what they are.

- Image n.2 is the “polygon annotation”, used for delineating irregular shapes (e.g. crosswalks), where the labeler must precisely define the irregular contours of different objects.
- Image n.3 is called “polyline annotation” and is used for lane detection (e.g. double lane, broken lane).
- Image n.4 is the “3D cuboid annotation”, specifically used for spatial measurements (e.g. cars length, height, depth).
- Image n.5 is called “semantic segmentation”, which allows to specify the image repartition by carefully outlining the various objects in the image (Anolytics, 2020).

All these labels serve as additional information, necessary to train machine-learning models to learn how to see and recognize different objects. These practices can be understood as infrastructural components where the scope, relevance and extension of the human infrastructure within the larger domain of AI, is proportional to the amount of work needed to make AI-based systems work through these practices. Navigating between *local* and *global* dimensions of scale gives an idea of how large this proportion is. In particular, analyzing the specific tasks of image labelling, firstly shows what is necessary for the advancement of autonomous cars to occur, and secondly, shows how humans are involved in it: if we think about the huge amount of objects and elements that need to be classified (considering also that each object needs thousands of labels), we can therefore imagine how large the human infrastructure is. To keep in mind, this example of labelling for autonomous cars industries can be extended to almost any imaginable area where these practices are applied: this may concern agriculture (e.g. crops health monitoring) where fields and lands are labeled, manufacturing (e.g. automated supply handling), security (e.g. human tracking) and many, many more (Anolytics, 2020).

AI Impersonation

After preparation, the second pole identified by Tubaro et al. (2020) is AI impersonation, which refers to the old concept of the Mechanical Turk. The idea of the Mechanical Turk comes from an eighteenth century fake chess-playing machine, which was apparently capable of playing chess on its own, while in fact it was secretly operated from the inside by a human player (p.7). Similarly, Tubaro et al. (2020) point out that some companies that claim to provide AI-based services, use instead human resources to partly (or entirely) provide those services: for example, a start-up called Julie Desk created to produce e-mail based scheduling assistants, worked for months by relying on humans instead of using, as declared, an automated algorithm; and even when the algorithm was finalized, humans continued to be involved in the loop to validate the final steps of the process before reaching out to customers (p.7). Another example, among many others, is reported in an article of The Guardian

(2018) in which Expensify, a business management application which was supposed to use "smartscan technology" to transcribe customers' receipts and documents, was instead relying on human workers to manually process and transcribe those texts, which were then passed off as technological outcomes. These examples illuminate some key aspects of how workers invisibility can occur, showing how capabilities that are generally attributed to AI, should in fact be attributed to the work of human operators.

AI Verification

The third pole identified by Tubaro et al. (2020) is verification and represents the last part of a larger system in which humans are invisibly involved. During the verification phase, humans are employed to verify the outputs of an AI automated system (p.9); the main goal is to check the quality and accuracy of the outcome, and eventually correct it. One of the most common examples of verification is content moderation, which in itself can be considered as an act of classification that, rather than happening during the preparation phase, occurs towards the ending part of the process to validate or deny the circulation of online materials. Content moderation is another example where human functions are often concealed by the overestimated capabilities of AI: in *Behind the screen: Content moderation in the sh adows of social media,* Roberts (2019) explains that despite the general belief that social media contents are supervised and secured by AI algorithms, "[...] much of the labor of these adjudication processes on platforms is undertaken not by sophisticated artificial intelligence and deep learning algorithms, but by poorly paid human beings who risk burnout, desensitization, and worse because of the nature of their work" (p.25).

This indicates that in order to maintain YouTube, Facebook and other digital platforms "clean", a lot of disturbing material (e.g., violence, child pornography, suicide) generated by users is checked, evaluated and then classified as adequate or not by an invisible human workforce. Because of the semantic and interpretative complexity involved in this process of classification, large part of social media contents requires human functions to be properly screened, although this is usually thought as being exclusively part of algorithmic work. While for a human operator it is relatively simple to understand the context and meaning of a digital content, for an automated system it is not; despite the stated progress of AI, there are still many aspects for which human functions are imperative, even if unseen.

The account provided by Tubaro et al. is useful to organize these practices, and to offer an organized picture of the various ways in which humans are involved in data processing. However, it does not entirely cover the multiple roles in which the human infrastructure branches out. For this reason, in the following figure I schematized some complementary key areas and activities in which humans operate in the context of AI:

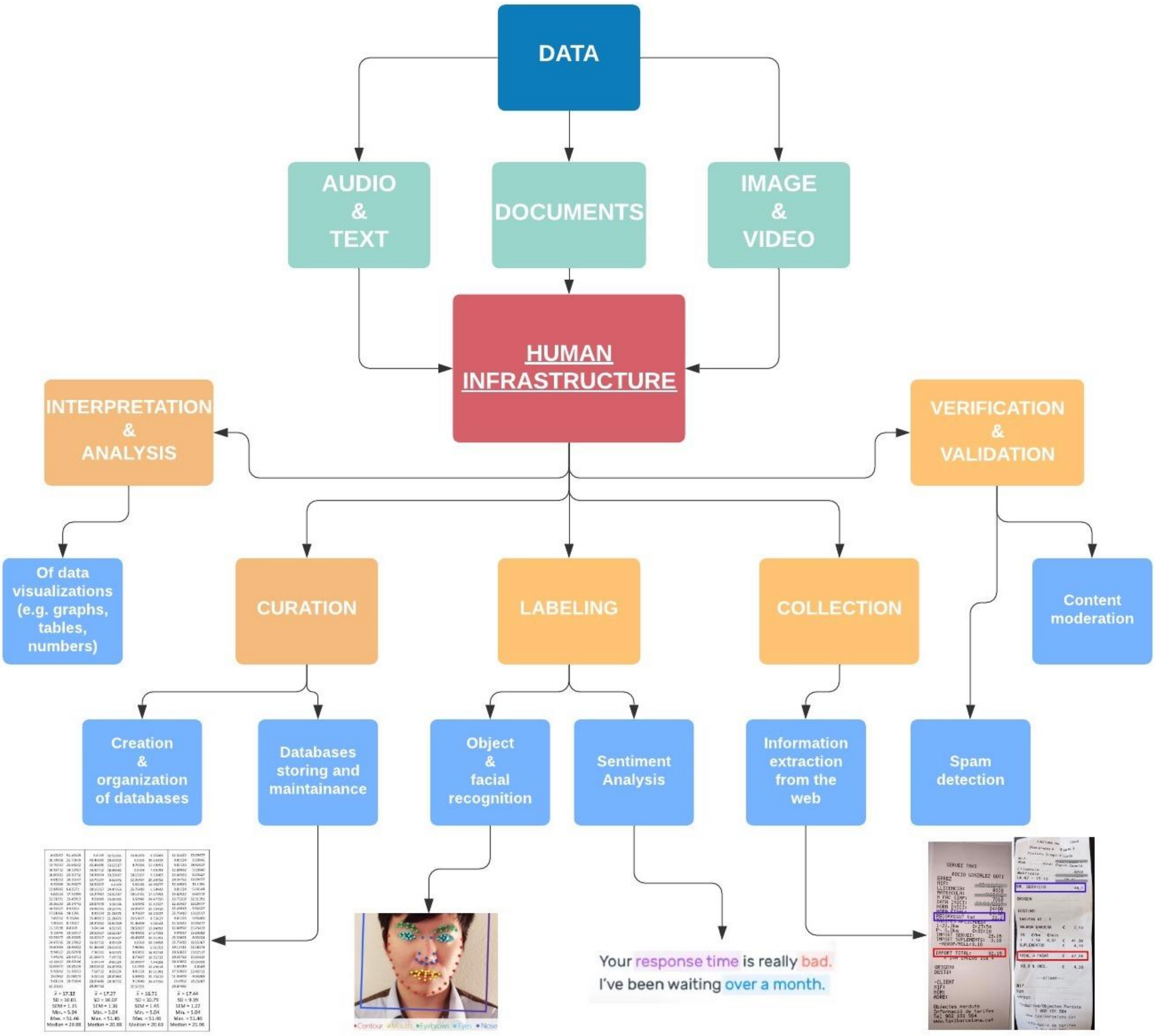

Fig.5 - Additional key areas and functions in which the human infrastructure is involved.

As can be noted from this scheme, humans are involved in many different ways, and as Tubaro et al. (2020) correctly suggest, the contribute that they provide to developing AI with these tasks is *structural* rather than *temporary*; there is no valid reason to think that these activities will disappear or become less relevant in the future. However, rather than as *structural*, thinking about humans and their contribution as *infra-structural* emphasizes on one side their status of necessity, solidity and durability; but on the other side, they are in this way framed as the locus in which the continuous changes, negotiations and tensions proper of infrastructures occur. As infrastructure, humans are entangled and embedded in various organizational infrastructures - like tech companies and service providers - and their work is often mediated by means of a networked variety of digital platforms, apps and other technologies. In the next section, I will elaborate on the forms of organization through which these working processes occur.

2.7 The organization of the Human Infrastructure

The various practices highlighted so far, represent the function that the human infrastructure occupies in the context of AI. However, these practices are organized, managed and performed across a complex system of interconnected infrastructures, of which the human is just one among many. In fact, one of the main characteristics of infrastructures is *embeddedness* (Star et al., 1996, 113); the human infrastructure is sunk into a thick network of technological and organizational infrastructures, that this section aims at disentangling. I will therefore expose the organizational processes and technological dynamics in which humans are situated to illustrate how the human infrastructure is structured.

If we imagine that the body of humans composing an infrastructure creates an overarching form, one may try to visualize how this form looks like: for example, infrastructures like railways, electricity networks, etc. usually have a reticular shape, which interconnected fabric is composed of a series of junctions between nodes and lines. Similarly, one can think about the human infrastructure as having a similar shape, of which nodes and lines branches off around the globe. At the same time, since the infrastructure is approached here as dynamic and emergent in its form, the contours of its shape are not easily definable (Harvey et al., 2016). Actually, it is better to imagine it without contours at all. Nodes and lines are helpful metaphorical tools to represent the various parts composing the infrastructure. Nodes, on one side, represent the most relevant organizational and technological clusters around and through which the practices highlighted until now unfold. These are, for example, specific digital platforms[8] to which workers access in order to label data; service companies that dispose of a workforce ready to perform certain tasks; factories of data where workers gather to carry out related activities. Lines, on the other hand, represent the connective fiber that links clusters to individuals, and individuals to clusters. They symbolize the transcendence of time and space, for which different geographical locations can be instantly reached and networked. For example, a labelling task can be performed on the same digital platform by a person living in India and by one living in Spain at the same time; but what matters, is that despite the geographical distance, the two are unified as part of the same human infrastructure by the connective shape of its fabric. Now, to

[8] A looser definition of a digital platform, is one in which social and economic interactions are mediated online. (Kenney et al., 2016) For the scope of this research, I will be mainly concerned with the working dimension that digital platforms allow to mediate, and how different forms of invisibility depend on it. While acknowledging the fact that digital platforms can mean different things, and are different from each other, I will primarily consider their shared capacity to "transcend" space and time, and accordingly shape working and societal dynamics.

make things more concrete, I will analyze four different types of working organization through which the human infrastructure expands, and by which its functions (e.g. labelling, verification etc.) are enabled.

In-house labour

The first type of organization is "*in-house labour*" and represents a more traditional way in which tasks and workflows are organized. Here, tasks are executed by regularly paid employees inside the company, who tend to provide high levels of accuracy in data processing (CloudFactory, 2019). However, a higher level of accuracy is undermined by slowness and the high amount of time that the process requires. In fact, the labour in house usually relies on a limited number of humans, and this becomes problematic when the amount of data to label scales up. At this point, the question of scale - which has been previously identified as the main tension from which the human infrastructure grew - becomes fundamental. The necessity to classify a huge amount of data is at odds with the scarce availability of labels which can be provided by a limited number of workers; this tension marks the expansion of the human infrastructure, from which the following forms of organization have taken shape. Importantly, the following forms of organization have not been instantly created out of nowhere, but, emerging from already existing infrastructures, they have been rather arranged together as results of adjustments and negotiations: since infrastructures are layered and complex, they are fixed in modular increments rather than globally or all at once (Star, 1999, 382).

Crowdsourcing

The second form of organization is "*crowdsourcing*", also defined as "the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined, and generally large, network of people" (Guittard et al., 2015, 50). This practice emerged as a response to the scale problem, thus allowing to obtain labeled data in a cheaper and quicker way; consequently, the quantity and availability of labeled data increased significantly (Lease, 2011). The organization of labour through crowdsourcing, relies principally on the use of digital platforms which support and enable the workflow, and at the same time replace some organizational aspects of work (Berg et al., 2018, 6). Digital platforms like Amazon Mechanical Turk, Appen, Lionbridge are easily accessible to people from around the world to work from their homes. In figurative terms, digital platforms can be imagined as the nodes around which human labour is clustered, and the lines as the connective fibers of the networked (infra)structure that links geographically distributed individuals to the digital platform. Differently from in-house labour, where the work is performed by specialized employees, crowdsourcing allows any individual having an internet connection to work on various tasks. For example, the Mechanical Turk of Amazon (2020), has its own “on-demand workforce” composed of 500,000 workers spread in 190 countries around the world, satisfying customers’ requests 24/7. Digital platforms facilitate new ways to commodify labour, that is sold to companies “on-demand” (Berg et al., 2018, 6). With crowdsourcing practices, the work is split into “micro-tasks”, namely small units of work that are quick to complete (e.g. recognizing multiple times the color of cars in various images), which are divided among the crowd of workers.

Various kind of tasks are performed on these platforms: some of them (e.g. data generation, correction, label, sort), are intermediary steps for the realization of machine-learning models. Some others (e.g. filling out questionnaires, online surveys, websites testing, content review and description) are instead not necessarily linked with AI data production or processing. In fact, the use of crowdsourcing for the generation and elaboration of data for AI systems is relatively recent and exemplifies how already existing technological infrastructures can be re-arranged to satisfy new needs.

Outsourced labour

The third form of organization is “*outsourced labour*”, in which the processing of data is outsourced to external parties. Companies have two main possibilities to outsource labour: to individuals or to other companies. When the labour is outsourced to individuals (e.g. temporary contractors and freelancers), these are generally hired through social media or job search websites. Differently from crowdsourcing, they are regularly screened and selected from the company, which manages the workflow and the payments. However, the problem of scale may persist if the amount of data to label is high.

The most adopted alternative as a response to the problem of scale, is outsourcing labour to other companies which are specialized in data processing (Cognilytica, 2019). In this case, the data processing is directly managed by a specialized company, that relies on its own workforce of employees or contractors. This allows them to act as intermediary between requesters and workers, guaranteeing higher levels of control and accuracy on the quality of labeled data. The workforce of one single company may count on thousands of workers, who can be managed in two different ways: in-house or remotely. When the workflow is managed in-house, outsourced companies usually dispose of physical spaces where workers can operate. These spaces may resemble “data labelling factories”, in which the organization of workers takes the form of a “cognitive assembly line”. The main difference with an industrial assembly line, is that the finalized product does not derive from the work of their hands, but rather from that of their minds.

Fig.6 - Workers inside the Infolks company building in Palakkad, India. (Infolks, 2020)

The pictures above display the workforce of an Indian based annotation company, consisting of a mid-size team of about 200 employees. On the other side, when labour is remotely managed, there are no physical spaces where workers operate, but the specialized outsourced companies can manage their own workforce to process data through digital interfaces. ScaleAI for example, is one of the most important emerging companies aiming at accelerating the development of AI applications. It has its own physical offices in San Francisco, where the main team of managers, engineers, etc. operates, but at the same time counts on a workforce of 30,000 contractors (ScaleAI, 2020), who are instead geographically displaced (e.g. Philippines) and work on data-related tasks through specific digital platforms. Similarly, there are many other companies that provide labelling services that, besides having a small team of employees, rely on a large number of disembodied digital workers living around the world to complete those tasks.

User-Driven

“User-driven” is the fourth way in which data for developing AI systems are obtained. As the name suggests, the process of data generation and processing derives from people using software and digital applications, or more generally from web surfers. The main idea behind a user-driven way of obtaining data for AI, is that users can become a resource to label data while making use of digital services. One of the most common examples is the Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHAs), namely security measures used on the web to prevent people from using computer programs to abuse of a service (e.g. ticket scalpers buying thousands of tickets online) (Law et al., 2011, 48). An example of CAPTCHAs is when a user must recognize and transcribe a piece of text correctly (which usually consists in 1 or 2 slightly distorted words) before gaining access to a website or digital service.

Although on one hand CAPTCHAs are applied as security measures, users are (often unknowingly) on the other hand participating in something bigger than that, named as the "largest distributed human collaboration project in history" (Law et al., 2011, 48). More than 200 million CAPTCHAs (which take 10 seconds of time on average) were typed every day in 2011, that means approximately 500,000 hours per day; the project reCAPTCHAs aims at redirecting this collective humans effort towards the optimization of AI text recognition systems within the massive wave of physical books digitalization (e.g. Google Books, Internet Archive) (Law et al., 2011, 48). Since the Optical Character Recognition (OCR) AI system that scans physical books to create digital texts is sometimes unable to recognize words (especially from ancient books with ink-faded words), human accuracy in recognizing those words is employed to improve the system (Law et al., 2011, 48). Similarly, the contemporary version of CAPTCHAs, which requires users to recognize different road objects (e.g. cars, trucks, road signs) in an image divided in 9 blocks, is likely to improve computer vision for self-driving cars industries by drawing on the aggregation of human collective inputs.

Another example that shows how users are intertwined with the production and labelling of data for AI systems are the "human-computation games", or "games with a purpose" (GWAP). The GWAP are online multiplayer games wherein data are generated as a byproduct of play (Law et al., 2011, 63). The players from which those data are obtained, are often unaware of the fact that the accomplishment of specific tasks lurks behind their playful activities (Bozzon et al., 2016, 423). The gamification of data labelling may disguise itself as a simple matching games, where for example the same image is given to two players, who are asked to describe it with a tag. If the image gets successfully tagged by both players with the same tag, they are rewarded with bonus points (Law et al., 2011, 63). This gaming mechanism acts as a double evidence, in which having two players sticking the same tag provides the system with a higher level of confidence about the correctness of the tag than a single player could provide (Law et al., 2011, 63).

To conclude, one can imagine that of humans as an infrastructure in which multiple organizational and technological layers intersect with each other. In order to specify how these various layers are arranged together, I divided four different forms of organization (in-house, crowdsourcing, outsourced and user-driven) which organically contribute to the global development of AI in different ways. Although creating this categorical division is important to understand more in detail the dynamics according to which data are generated and labeled, what is even more important is that these different forms of organization refer to the same infrastructure of humans. In fact, only when all these scattered pieces are considered as a whole, is possible to see how great the scope and relevance of humans within the world of AI is. After having described more specifically the practices and forms of organization in which the human infrastructure ramifies, the next section will elaborate on how humans and their labour are kept invisible, thus refining and expanding the concept of *invisibility* in relation to the human infrastructure.

Chapter 3
Infrastructural Invisibilities

Now that I have provided a more nuanced perspective on what the practices and forms of organization at work are, I will delve deeper into the notion of *invisibility* concerning the human infrastructure of AI, to show how the question of definitions as introduced in chapter 1, and the focus on the human infrastructure as described in chapter 2, help to illustrate in more detail how workers invisibility occurs in the context of AI. I will therefore describe the mechanisms that render and keep the human infrastructure invisible from view, by focusing on the question of definitions and the role of digital platforms. This will be done to show how invisibility is not simply a property of infrastructures but is also a complex mechanism of concealment that emerges in relation to specific practices and uses.

3.1 Invisible Labour

The emergence of new ways of working due to the development of information and communication technologies, has been researched across a colored variety of disciplines, from social and cultural studies to the study of new media and digital platforms (Fuchs, 2014; Huws, 2014; Scholz, 2012). The notion of "digital labour" summarizes in itself the study of the dynamics that arise from the intersection between digital technologies and work. Digital labour has been object of sociological, law and political studies aiming at addressing and problematizing the related phenomena of invisible labour. The investigation of invisible human labour, in turn, draws on a line of research begun with the work of a number of feminist scholars, who problematized several activities taken by female workers that, despite their being essential, were not formally recognized or compensated as work (Whittaker et al., 2018, 35). Studies on digital and invisible labour naturally conflated with the studies on AI, due to the amount of unrecognized and undercompensated work involved in the digital economy, inextricably connected to the development of AI systems.

In *Invisible labor: Hidden work in the contemporary world*, Crain et al. (2016) reflect on the reason why some kinds of work are more visible than others, and on what are the forces and trends preventing people from "seeing" the work; they want to understand what Hatton (2017), in a similar way, defines as "mechanisms of invisibility". By drawing on both sides' insights about workers invisibility, I will illustrate the processes that render and keep the human infrastructure of AI invisible from view, relating those intuitions to what has been formulated in chapter one and two. In the following paragraphs, I will illustrate two key factors, definitions and digital platforms, that contribute to make and keep workers invisible in the context of AI.

3.2 Definitions

The first factor that contributes to workers invisibility are the definitions. As claimed in chapter one, each definition can illuminate some aspects of AI while obscuring some others, portraying a specific picture of what AI is. As shown, in order to see the role that humans play in AI data production, labelling and processing, one must firstly understand that the AI paradigm involves huge amounts of data, and consequently that humans play a central role in their processing. The problem, is that definitions can hinder this understanding: when AI is simply defined as a technological application; when loose or no definition is provided; when AI capabilities are over-estimated and polarized; in all these cases, definitions lead us astray from a way of thinking that allows to see the deeper social implications involved, and therefore the invisible work of the people implied. In fact, only when the relevance of data and their processing is made explicit, one can start to see how and why workers invisibility takes place. In order to literally see invisible work, one needs to know how invisibility articulates through working practices of labelling, for instance. But in order to know how these practices occur, there should be a way of thinking about AI that allows to do that, which definitions can favor or hinder. This is particularly important if we consider for whom definitions make the human infrastructure invisible: in a survey conducted by Weber Shandwick and KRC Research (2016) across the globe (U.S, Canada, UK, China and Brazil), participants were asked where their conception of AI came from, and 80% of respondents mentioned a form of media, ranging from Internet and social media to TV and news articles or reports (Weber & KRC Research, 2016, 7). How media define AI, plays therefore a central role in favoring or hindering the visibility of labour for people whose conception about AI derives from media themselves.

3.3 Digital platforms

The second key factor that contributes to make and keep workers invisible, lies in the use of digital platforms in the organization of labour. Although invisible labour is not a new phenomenon, the Internet and communications technology have created new forms of invisible work, in which workers and work are hidden from view (Crain et al., 2016, 71). Among the various forms of organization identified so far, in-house labour is the one that is less affected by the mediation of work through digital platforms; in fact, since workers have a physical space where they can operate, digital platforms are not key elements to carry on the work. It is indeed the only form of organization that resembles more visible forms of work, according to the definition of Crian et al. (2016). They state that *visible labour* has “traditionally been defined as work that is readily identifiable and overt. It is located in a physical workplace and is self-recognized as work by management, employees, and consumers. It is typically paid, occurs in the public sphere, is directly profit generating, [...]” (p.3). Although the concept of visible work has naturally changed over time, digital platforms contribute to erode the definition of visible labour, enabling what Hatton (2017) calls “sociospatial mechanisms” of invisibility: invisible work is devalued because it is physically segregated from the “workplace”, and is performed in the domestic sphere or other non-traditional worksites (p.343). Accordingly, “Because of this spatial dispersion, digital workers are isolated from each other as well as from the businesses for which they work” (Hatton, 2017, 344). Workers are therefore less likely to organize, to ask for public support, or to appeal to the sociolegal system to improve their working conditions (Crain et al., 2016, 5). Sociospatial mechanisms of invisible work refer especially to the organizational forms of *crowdsourcing* and *remote outsourced* labour, which rely on the use of digital platforms to function. In the case of crowdsourcing, digital platforms are known as crowdsourcing platforms (e.g. Amazon Mechanical Turk, Appen, Lionbridge), in which people from all over the world can easily access, sign in, and start to complete their labelling tasks. In the case of remote outsourcing, digital platforms are referred to as Application Programming Interfaces (API), namely digital interfaces which connect and share data with companies that provide labelling tasks among geographically dispersed workers.

In both cases, the kind of work enabled by the use of digital platforms exacerbates labour fragmentation, of which dispersed pieces are harder to see, both visually and symbolically. Visually, because the visual act of seeing workers is missing. Symbolically, because often tasks are categorized as “not work” (Crian et al., 2016, 6). This symbolic aspect is compounded by what Poutanen et al. (2019) call *taskification*, namely the idea that through digital platforms, working activities can be sliced in smaller pieces (hence, tasks) which can be dissected from what is more traditionally thought of as work or professional work. If one thinks about the role of crowdsourcers in Amazon Mechanical Turk, it is indeed really hard to firstly see the workers, and to secondly identify whether their status can be categorized as that of a worker or not. A detriment to more traditional forms of labour (in-house), the forms of organization based on the use of digital platforms (crowdsourcing, remote outsourcing) are however better responses to the problem of scale. In fact, they are more and more adopted, allowing to defeat spatial and temporal distances, thus embracing unlimited resources of people, available 24/7, from all over the world.

3.4 Invisibilities emerging from users' interaction with digital platforms

There are two other main ways in which workers invisibility can manifest itself, and is constituted from users' interaction with digital platforms: in the first case, invisibility occurs between the user of a digital platform and the workers operating behind it, which renders workers' labour invisible to the user. The second, as in the *user-driven* case, occurs between users and use: in this instance, users make use of digital platforms, while unknowingly producing "work" that is invisible to users themselves.

In the first case, as mentioned in chapter 2, humans *impersonate* AI. The basic idea behind impersonation, is that the skills that should be attributed to the work of humans, are instead attributed to AI. The question should be asked, however, who attributes these skills to AI rather than to human workers – that is, the user. Users of digital platforms may not even know that people, rather than programs or algorithms, are actually working (Crain et al., 2016, 73). For example, in the case of content moderation, the user of a social media like Facebook could think that feed cleaning is due to smart algorithms that automatically eliminate inappropriate contents, while it is also because of humans screening that violent or discriminatory contents are removed. An illustrative example of workers invisibility due to the interaction with digital platforms is given by Taylor (2018) in this short anecdote:

One recent afternoon I stood waiting at a restaurant for a to-go meal that I had ordered the old-fashioned way – by talking to a woman behind the counter and giving her paper money. As I waited for my lunch to be prepared, the man in front of me appeared astonished to receive his food. "How did the app know my order would be ready twenty minutes early?" he marveled, clutching his phone. "Because that was actually me," the server said. "I sent you a message when it was done."

This anecdote describes the mechanism of invisibility that contributes to keep workers in the background, hidden from the view of the user due to the technological interaction which lies in between. At the same time, users tend to attribute automation capacities to technologies that are visible in the foreground. In the second case, the mechanism of invisibility does not occur between users and workers, but between users and use. Or, in other words, between users and users themselves. Here, the work is rendered invisible before the eyes of the user that is actually performing it. An example of this invisibility can be found in the concept of *gamification*. Gamification represents the idea that the line that used to separate work and leisure, is blurred by means of technology;

gamification through digital platforms hides work as leisure, obscuring what in more traditional circumstances would constitute an employer-employee relationship (Crain et al., 2016, 73). As in the case of the GWAP mentioned in chapter two, gamification includes some forms of work that do not feel like work to the user performing the task (Crain et al., 2016, 80). They can instead be funny, but at the same time aiming at solving AI problems that, more traditionally, would require a worker to be accomplished. In the following image, some examples of GWAP with the related AI problems that they aim at solving are illustrated:

Game	Description	Mechanism	AI Problem
The ESP Game [334]	two players match on a tag for the same image	output-agreement	object recognition
Peekaboom [337]	player 1 reveals parts of the image associated with a secret word, player 2 must guess the secret word	function computation (problem inversion)	object location
Verbosity [338]	player 1 describes the properties of the entity associated with a secret word, player 2 must guess the secret word	function computation (problem inversion)	knowledge extraction
TagATune [190]	two players exchange tags and determine if two music clips are the same or different	function computation (input agreement)	music classification
FoldIt [69]	players fold protein structures to minimize total energy	function computation (optimization)	protein folding
HerdIt [24]	players select tags that describe the music	output-agreement	music classification
Categorilla [329]	players name an entity that fits a template (e.g., Things that fly)	output-agreement	natural language processing
MoodSwings [174]	players click on a 2-dimensional grid to indicate the valence and intensity of the mood of a music clip	output-agreement	music mood classification
Phrase Detective [57]	players identify relationships between words and phrases in a short piece of text	output-agreement	natural language processing
Phylo [255]	players align colored blocks by moving them horizontally and inserting gaps	output-agreement	genome alignment

Fig.7 - Several examples of games with a purpose (Law et al., 2011, 69).

As one can see from the figure above, there are many games adapted for many AI problems, to which users can unknowingly provide solutions by playing. Much of the invisible work naturalized as part of what it means to be a user in the digital age, is indeed not recognized as economically valuable as it actually is (Ekbia et al., 2017). These cases show that the human infrastructure of AI is configured in such a way that allows various kinds of invisibility to manifest, which may vary in relation to context and use. When analyzing larger systems, framing invisibility as if it were just a given property of infrastructures means to simplify it and failing to see its nuances. It is therefore more effective to conceptualize invisibility as a relational property of infrastructures that arises through interactions, and that may vary according to different contexts and uses: “For a railroad engineer, the rails are not infrastructure but topic” (Star, 1999, 380). In this sense, one may wonder what renders the human infrastructure invisible, and for whom. Rather than about invisibility, one could therefore think in terms of invisibilities.

Chapter 4
Ethical concerns on labour conditions and epistemological issues

But what does the conceptualization of the human infrastructure of AI as such - and the exposure of the practices, forms of organization and interrelations with technology - add to the debate about AI? To look closely at the human infrastructure means to recognize that many aspects of developing AI systems depend on a vast network of humans, who are inextricably involved in the process of data labelling and curation. To bring the human infrastructure out from the realm of invisibility allows to firstly raise and discuss ethical concerns about workers' labour conditions. Secondly, it enables epistemological claims against misrepresentative portrayals of AI to be made, thus showing how the infrastructural approach adopted so far enriches the debate about AI in regard to the social implications involved.

4.1 Labour Conditions

Despite a lot of research on AI is focused on the anticipation of potential scenarios, in which the impact of automated technologies will disrupt the job market and shape social structures in the future, less attention is put on how developing AI systems is already changing working and societal dynamics in the present. Unregulated and invisible forms of work, underpayment and issues of social security are already eroding the market of labour and transforming society. The literature on invisible labour has problematized the dehumanizing nature of practices of data labelling and curation, especially in relation to the underregulated and underpaid types of work in which they manifest themselves (Hara et al., 2018; Berg et al., 2018). This has naturally converged with the social and political issues that have been covered in the literature on platform-work (Graham et al., 2017; Choudary, 2018; Berg, 2016). In this regard, several ethical concerns have repeatedly emerged, especially in relation to the poor working conditions to which workers are subjected.

In the case of work on crowdsourcing platforms, the practices of data labelling and curation are often described as labor-intensive and repetitive (Barbosa et al., 2019; Tubaro et al., 2019); more importantly, labor-intensiveness and repetitiveness are inseparable consequences of the underpayment inherent to these practices. A research covering 3,500 workers from different 75 countries on 5 different crowdsourcing platforms, found that a worker earns on average US$4.43 per hour (only when paid work is considered), while only US$3.31 per hour when also unpaid work is considered (Berg et al., 2018, 49); average earnings for paid and unpaid work varies from region (US$2.22 Asia and the Pacific) to region (US$1.13 in Africa) (Berg et al., 2018, 52). Given that workers are paid for each individual task they complete (starting from the minimum of US$0.01 per task), it follows that to earn some money (often far from the minimum hourly wage), the kind of work required becomes highly intensive and repetitive, thus dehumanizing in its very nature. Worryingly, for about 32 per cent of the workers in question, crowdsourcing work is the primary source of income; but, since crowdsourcing platforms categorize their workers as self-employed, they lack any kind of protection of labour and social security law (Berg et al., 2018).

There are also issues of power imbalance, due to the phenomenon of taskification: as can be read from Amazon (2020): “Pay only for satisfactory work – You do not pay a Worker or Mechanical Turk fees until you accept the Worker's work. [...] Workers don't get paid if you reject their work”. The division of work into tasks facilitates their own acceptance or rejection by the requester. In the case of rejection, the worker simply does not receive any payment. Being the working relation mediated by digital platforms, any disagreement or divergence about the rejection is hardly solved, thus contributing to the imbalance of power which exacerbates the status of the worker. These transformations of work arrangements are not new or unique but are part of a broader shift towards more precarious forms of labour, and more automated management processes (Berg et al., 2018, 6). They can ultimately be seen as “a throwback to the de-skilled industrial processes associated with Taylor, but without the loyalty and job security” (Cherry, 2016, 3). Even when the organization of labour seems to be more ethics-oriented, we should question it. In the case of some ourtsourcing companies, like the San Francisco-based company Samasource, thousands of workers are employed in developing countries; Samasource leads working centers in Kenya, Uganda, Nairobi and other regions, where they claim to drive social impact providing “dignified digital work and paying living wages” (Samasource, 2020). While it is undeniable that they contribute to improve poor people lives (paying 8-9$ those workers of the informal economy who usually earn 2$ or less per day), we should also consider that they outsource data labelling tasks for clients like Google, Microsoft, Walmart and Ford. These big companies can therefore profit from cheap labour while claiming to sustain poverty eradication one job at the time. A living wage of 8-9$, despite being a good paycheck in African countries, has a completely different value if situated in a global context, where large multinational companies can save millions of dollars by outsourcing data-related work in developing countries.

A five-years study conducted on machine-learning work in the African economy, has also highlighted how outsourced forms of work alienate workers from their own labour; since they do not know anything about the end-client or the application of the final product, many workers can only speculate about the utility of their jobs. When asked about the purpose of his job, as the study reports, a worker replied by saying “They don’t tell me; they just want lots of tagged images.” (Anwar et al., 2020, 5). I believe that an infrastructural approach like the one adopted in this research, can benefit the understanding of how these ethical issues arise in the context of AI. It can illuminate the dark corners where most of the invisible work underlying AI systems occurs, and is particularly useful to trace the complex systems of relations that link automated technologies to the hidden layers where humans operate to make them function: as shown in the previous chapters, autonomous vehicles are, for instance, directly dependent on how humans label images and video of streets and pedestrians. Thus, uncovering this interdependency opens the possibility to analyze in what the act of labelling consists of. One can therefore discover that the kind of work required from humans is generally shaped around the request to label, correct, structure, categorize or verify data, and has some recurring features. It is indeed highly standardized, repetitive and systematic in most of its manifestations, and this strongly limits the space in which human expression can flourish. Paradoxically, humans undergo machines by becoming mechanical, in order to teach machines how to become human. I am convinced that putting emphasis on such issues and raising this kind of awareness, can shine a light on those workers whose precariousness is also due to a lack of visibility that, as argued, contributes to deteriorate their status. As Irani (2016) nicely puts it, one should ask “What would computer science look like if it did not see human-algorithmic partnerships as an embarrassment, but rather as an ethical project where the humans were as, or even more, important than the algorithms? What would it look like if artificial intelligence and human-computer interaction put the human care and feeding of computing at the center rather than hiding it in the shadows?” (p.37).

4.2 Conceptual and epistemological considerations

Focusing on the human infrastructure also allows to question the consistency of the data-driven paradigm of AI, which envisions a future where humans will supposedly be less and less relevant due to the higher levels of automation that AI technologies are progressively capable of achieving. By exposing the processes that fall invisible behind this narrative of AI, some critical reflections to counter such accounts can be made on a conceptual level. It has been previously argued that public perception is influenced by how media define AI. Portrayals of AI in the public sphere are often characterized by an over-estimation of the automation capacities of artificially intelligent technologies, thus resulting in an artifactual description of what AI technologies can do. Such accounts are filled with expectations about the almost unlimited possibilities that technological progress opens to humans, without considering that, in reality, it is humans themselves who are actualizing those possibilities. Bringing the human infrastructure out from the realm of invisibility allows to see that the work which is wrongly perceived as automated, is more rightfully attributable to the manual and cognitive abilities of humans. More discussion should therefore be concerned with the human conditions within this process, in order to tackle the ethical challenges that the practices and the forms of organization highlighted so far bring to the table. At the same time, making the human infrastructure visible can provide the public with a truer picture of AI. In this sense, one can ultimately see how much of human is obscured by the notion of *artificial*, and how many repetitive, standardized, and trivial tasks are obscured by the notion of *intelligence*. Accordingly, this awareness can have demystifying effects on how the public perceive AI: rather than blindly accepting that a "useless class" (Harari, 2017) will inevitably emerge as a by-product of AI progress, it would be possible to realize that humans are deeply involved in those systems that the author predicts will occupy people's social and economic relevance. This enables a critical reflection about humans' own value and significance in the context of AI: it is indeed quite incorrect to say that AI systems will simply replace workers, thus leading to their exclusion and making them irrelevant. Contrarily, AI systems are extremely inclusive, in the sense that they involve a vast network of humans, who are instead utterly central in their process of development. Overturning the classical conception of AI as opposed, contrasting and antithetical to humans, and recognizing that humans actually form, constitute and shape AI, is of primary importance to escape the man/machine, human intelligence/artificial intelligence dichotomy.

Furthermore, there are epistemological aspects related to bias and artificial objectivity that a focus on the human infrastructure helps to consider. By showing how human cognitive functions are involved in developing AI systems, there are indeed interesting epistemological question to reflect on: first of all, looking at the practices in which humans are involved in developing AI systems, shines a light on the mechanisms that occur in the phase of data labelling that precede machine-learning models training. Data labelling is just one among many stages in the development and application of machine-learning models, but the potential biases introduced in the phase of data labelling remains largely unexplored (Barbosa et al., 2019). An infrastructural approach that helps to conceptualize the different ways in which humans label data, is therefore epistemologically valuable for assessing issues of bias and algorithmic opacity in the first stages of their development. For example, various authors have investigated how bias are introduced during the phase of data labelling through crowdsourcing (Barbosa et al., 2019; Eickhoff, 2018). I consider that the methods adopted in this thesis are particularly valid to expand these lines of research towards an infrastructural approach on humans and AI, thus contributing to evaluate more ways in which human labelling could introduce bias in the application of AI systems.

Moreover, as argued in Boon (2020), there are epistemologies from the empiricist tradition in the philosophy of science that support claims about AI objectivity. According to such views, algorithmic decision-making processes and predictions based on data are assumed to be objective in their outcomes. But as Leonelli (2016) argues, views on the objectivity of data hold if we consider data to be *given*, rather than *made*; the idea of having data as objective, context-independent sources of evidence, clashes with the complex processes through which they are obtained, manipulated and circulated, and for which a lot of human subjective judgements are involved. Views on AI objectivity are therefore at odds with the high volume of subjective, human-made decisions involved in the phase of data labelling as illustrated in chapter two. For instance, labelling tasks where subjective human judgements are required, are impacted by multiple workers demographics, belonging to subgroups, cultural, linguistic and cognitive differences, and many other significant variables (Barbosa et al., 2019) which challenge the ideal of AI as objective in its outputs. A focus on the human infrastructure can therefore illuminate these dark corners that a lack of attention in the public debate contributes to keep away from view. The question of bias concerning AI-based systems is tightly connected to the ideal of AI objectivity. Taken together, they seem to suggest that once that the problem of human bias is somehow solved, AI objectivity in the decision-making process will be possible. However, rather than asking whether it will be possible to achieve an artificial objectivity or not, one should look at how many human-made decisions are made along the process, and then draw his own conclusion. Analyzing the human infrastructure suggests indeed that as long as humans will be involved in-the-loop, the exclusion of bias and the achievement of AI objectivity are far from being achievable. Contrarily, since biases exist and are deeply rooted in what it means to be human, it is by recognizing that the ideal of AI objectivity depends on those biases that more meaningful ways of addressing this problem can arise. Rather than maintaining these practices invisible, a new approach would consist in making them extremely visible: by thoroughly looking at how human-made decision are made, in what the labeling tasks consist of, how these tasks are organized and made accessible, and many other strategies would indeed provide a clearer account of how human biases arise, and how to tackle them. However, this would mean to open the “black box” in which humans find themselves as an infrastructure, thus uncovering the whole set of ethical issues deriving from underpayment, poor labour conditions and scarce regulation, and perhaps the price to pay for it would be too high.

Conclusion

During the various conversations I have had on this research topic, a strong objection has often been raised against the view on humans and on their role within the AI system proposed in this thesis. The objection says that the infrastructure of humans will not be relevant in the next future, since AI will learn how to perform also those tasks for which humans are required today; accordingly, analyzing and mapping the role of humans and their labour in the context of AI will be irrelevant in the years to come, since they will have no role at all. This objection rightfully considers the increasing capacity of AI to perform human-like activities, that consequently includes also labelling tasks. In fact, the labelling paradigm is already evolving in terms of efficiency and sophistication, such as in the case of the so-called "Active Learning" model, which integrates AI functions in the labelling phase (Prendky, 2019). In this way, the labelling process is hybridized, by conjugating human capabilities with AI functions to optimize data labelling timing and accuracy.

Despite this objection implies that humans will become superfluous, there are good reasons to think that this will not be the case, and therefore counter it. Whether human jobs will be *replaced* by AI or not, is indeed the wrong way of framing the problem; trying to understand how jobs are rather *reconfigured* by AI technologies, leads instead to more meaningful ways of addressing the issue (Mateescu et al., 2019, 13). This is what Gray et al. (2019) call the "paradox of automations' last mile", namely the idea that since the boundaries of automation will always be pushed one step further, there will always be some new configuration of humans that will be arranged to complete new tasks; the "last mile" represents in fact the gap between what a human can do and what a machine can do. ImageNet, a rich dataset of human-labeled images used to develop sophisticated image recognition algorithms, is one example of how this paradox occurs. ImageNet allowed to accelerate the development of AI systems to recognize images; however, even when AI technologies became able to recognize images (assuming here that they did it correctly, and better than humans), that would not mean to replace humans in recognizing images. In fact, researchers then focused on how to teach AI how to recognize *where* an object is situated in an image, for which more training data and therefore humans were needed (Gray et al., 2019). Human labour was not replaced, but rather reconfigured in a different way as a result of technological advance. Accordingly, analyzing how the human infrastructure evolves in the context of AI is not only important, but necessary to understand how its continuous unfolding forms of organization and practices dialogue with technological progress. This also allows to identify and discuss issues of ethical and epistemological relevance, which can be more easily intercepted by adopting an infrastructural approach that considers how various technological advancements are dependent on hidden working practices. The act of tracing relationships and interdependencies can therefore teach how to distinguish between jobs disappearing and jobs invisibility, acknowledging that the latter may still hide forms of human labour.

Finally, some limitations of the research can be identified. The first limitation regards the lack of a distinct focus on the technologies cited in the previous chapters. The use of digital platforms, as a consequence of the focus on larger infrastructures rather than single technologies, is loosely framed as part of the technological layers in junction with the infrastructure of humans with which they interact; however, an in-depth case analysis of how individual digital platforms shape specific working dynamics would enlarge and benefit the scope of the research. This could be done in a similar way to how Kate Crawford and Vladan Joler (2018), in their visual essay A*natomy of an AI system*, made visible and explicit the whole infrastructural network of materials, workers, organizations, technologies, knowledge, environmental resources, etc. needed to build a single Amazon Echo, showing how complex the system of connections and interdependencies is. Secondly, there are many others features of infrastructures, as described from various authors in the field of Critical Infrastructure Studies, which may be used to expand the conceptual boundaries that limit the notion of human infrastructure of AI in this thesis; here, I have only considered some of them to avoid resulting too general or vague. However, considering and including also other dimensions of infrastructures would enrich the object of this thesis and widen its horizon.

The more complex and interrelated the evolving system of interactions in the context of AI becomes, the more an infrastructural way of thinking may benefit the understanding of the dynamics which lie at the intersection between humans and AI. In this thesis, I have adopted an infrastructural approach to answer the research question "What is the picture of AI that emerges when it is analyzed through the lens of the human infrastructure that underlies its development?". This question implies that even before starting to provide an answer, it first needs to be clear what is meant with the terms "AI" and "human infrastructure", and then how these two dimensions come together in the same research question. In the first chapter I have therefore drawn on the literature on AI and machine-learning to shown that the act of defining itself is already crucial for framing AI in a certain way. I have illustrated that media can influence the public perception in a way that highlights the automation capacities attributed to AI systems, thus introducing the ideological function that tends to oversell automation. I have then provided a relevant definition of AI for the scope of this thesis, which allowed to analyze the basic structure of machine-learning, thus introducing the role of the humans in the context of AI. In this way, I have suggested that what is usually thought to be automated and disruptive, involves instead the work of many humans. In have then started the second chapter by expanding on the role of humans, by conceptualizing them as an infrastructure. I built my analysis on the theory and methods of the field of Critical Infrastructure Studies (STS), focusing on the specific infrastructural properties of *invisibility*, *embeddedness*, *reach*, *scope* and *scale* and the method of "infrastructural inversion". By combining these theoretical elements of infrastructures with this methodological approach, I have brought to the foreground the human network of arrangements, practices and forms of organization that constitute the human infrastructure, to show how humans are deeply involved in practice in the process of developing AI systems through tasks of data labeling and curation. Providing a more nuanced perspective on what the practices and forms of organization at work are, allowed me to analyze more in detail the property of *invisibility* concerning the human infrastructure of AI. Accordingly, in the third chapter I related to the question of definitions and the focus on the human infrastructure provided in the previous chapters. In particular, I have focused on the role of definitions and digital platforms to show how workers invisibility occurs in the context of AI, thus providing a clearer description of how invisibility articulates through working practices of data labelling and curation. After having exposed what lies in the intersection between humans, data and digital platforms, in the last chapter I focused on the ethical, conceptual and epistemological issues that making the human infrastructure visible allows to discuss. Basing on what has been formulated in the previous chapters, I raised ethical concerns about the marginalized role of workers in developing AI system and their labour condition; then, I reflected on how a perspective on AI through the lens of the human infrastructure can illuminate new ways of conceptualizing it and can challenge accounts which describe AI as automated and disruptive; finally, I related the information provided on data work to the epistemological issues concerning bias and AI objectivity, to highlight how an infrastructural approach to AI could open new ways of addressing such issues.

Fig.8 - "Snap#2" by artist Bruce Gray.

In the end, the picture of AI that emerges when it is analyzed through the lens of the human infrastructure that underlies its development is a complex one, composed of many patches of colors mixed with each other, and interspersed with white straight lines on a black background. It is definitely more complex than a single definition of what AI could mean, and it is certainly more difficult to grasp than a description in words can do in this thesis. It is just one of the many pictures that can represent the development of AI through the lens of the human infrastructure and can be interpreted from person to person in a different way. My position in this regard, is that the moving, colored stripes representing the network of humans seem to belong to the *background*, which is limited by the order dictated from the straight white lines representing AI and the development of its system *in the foreground*. Despite there seems to be a contrast between back and front, if one stares for a few seconds on a fixed point of the image, he will start to see that all the lines begin to blend into a single picture, where the colored and white lines become indistinguishable from each other. Then, one can notice that there is no clear distinction between the two dimensions, and that suggests that the human/AI dichotomy previously identified as such, does not really exist. What rather exists, is a way of concealing how humans are involved in developing AI systems, and many other ways of telling how AI technologies can do without humans. In order to counter such views, I provided an analysis of AI through the lens of the human infrastructure to show how humans shape AI and, in doing so, I hope that I made humans a little more visible than before.

References

Abu-Mostafa, Y. S., Magdon-Ismail, M., & Lin, H. T. (2012). *Learning from data (Vol. 4).* New York, NY, USA: AMLBook.

Alpaydin, E. (2010). *Introduction to Machine-learning Second Edition*. MIT Press.

Anwar, M. A., & Graham, M. (2020). *Digital labour at economic margins: African workers and the global information economy*. Review of African Political Economy, 1-11.

Barbosa, N. M., & Chen, M. (2019). *Rehumanized crowdsourcing: a labelling framework addressing bias and ethics in machine-learning*. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-12.

Barrow, H. (1996). *Connectionism and neural networks*. In Artificial Intelligence 135-155. Academic Press.

Bereiter, C. (1991). *Implications of connectionism for thinking about rules*. Educational researcher, 20(3), 10-16.

Berg, J. (2016). *Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers*, in Comparative Labor Law and Policy Journal, Vol. 37, No. 3, 543–576.

Berg, J., Furrer, M., Harmon, E., Rani, U., & Silberman, M. S. (2018). *Digital labour platforms and the future of work: Towards decent work in the online world*. Geneva: International Labour Office.

Boon, M. (2020). *How Scientists Are Brought Back into Science—The Error of Empiricism*. In A Critical Reflection on Automated Science. Springer, Cham, 43-65.

Bostrom, N., & Yudkowsky, E. (2014). *The ethics of artificial intelligence*. The Cambridge handbook of artificial intelligence, 1, 316-334.

Bowker, G. C., & Star, S. L. (1998). *Building information infrastructures for social worlds - The role of classifications and standards.* In Community computing and support systems, 231-248. Springer, Berlin, Heidelberg.

Bowker, G. C., & Star, S. L. (1999). *Sorting things out* (Vol. 297). Cambridge, MA: MIT Press.

Bozzon, A., Cudré-Mauroux, P., & Pautasso, C. (Eds.). (2016). *Web Engineering: 16th International Conferenc*e, ICWE 2016, Lugano, Switzerland, June 6-9, 2016. Proceedings (Vol. 9671). Springer.

Brennen, J., Howard, P.N & Nielsen, R.K. (2018). *An industry-led debate: How UK media cover artificial intelligence*. Reuters Institute.

Castells, M. (1996). *The Rise of the Network Society*. Cambridge, MA: Blackwell Publishers

Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018*). Portrayals and perceptions of AI and why they matter*, Royal Society.

Cherry, M.A (2016). *Beyond misclassification: The digital transformation of work*. Comparative Labor Law and Policy Journal, Vol. 37, No. 3, 544–577.

Chew, R., Wenger, M., Kery, C., Nance, J., Richards, K., Hadley, E., & Baumgartner, P. (2019). *SMART: An Open Source Data Labelling Platform for Supervised Learning. Journal of Machine-learning Research,* 20(82), 1-5.

Choudary, S. P. (2018). *The architecture of digital labour platforms: Policy recommendations on platform design for worker well-being*. ILO Future of Work Research Paper. International Labour Organization.

Chuan, C. H., Tsai, W. H. S., & Cho, S. Y. (2019). *Framing Artificial Intelligence in American Newspapers.* In Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 339-344.

Ciborra, C. U., & Hanseth, O. (1998). *From tool to Gestell.* Information Technology & People.

CloudFactory (2019). *The Ultimate Guide to Data Labelling for Machine-learning,* Report.

Cognilytica (2019). *Data Engineering, Preparation, and Labelling for AI 2019*, Report.

Crain, M., Poster, W., & Cherry, M. (2016). *Invisible labor: Hidden work in the contemporary world.* Univ of California Press.

David, R. A., & Nielsen, P. (2016). *Defense science board summer study on autonomy*. Defense Science Board Washington United States.

Dubljevic, V. (2012). *Principles of justice as the basis for public policy on psychopharmacological cognitive enhancement.* Law, Innovation and Technology, 4(1), 67-83.

Edwards, P. N. (2003). *Infrastructure and modernity: Force, time, and social organization in the history of sociotechnical systems*. Modernity and technology, 1, 185-226.

Edwards, P. N. (2010). *A vast machine: Computer models, climate data, and the politics of global warming.* Mit Press.

Edwards, P. N., Bowker, G. C., Jackson, S. J., & Williams, R. (2009*). Introduction: an agenda for infrastructure studies.* Journal of the Association for Information Systems, 10(5), 6.

Eickhoff, C. (2018*). Cognitive biases in crowdsourcing. In Proceedings of the eleventh ACM international conference on web search and data mining*, 162-170.

Ekbia, H. R., & Nardi, B. A. (2017). *Heteromation, and other stories of computing and capitalism*. MIT Press.

Ensmenger, N. (2018). *The environmental history of computing*. Technology and culture, 59(4), S7-S33.

Fast, E., & Horvitz, E. (2017). *Long-term trends in the public perception of artificial intelligence*. In Thirty-First AAAI Conference on Artificial Intelligence.

Fuchs, C. (2014). *Digital Labour and Karl Marx*. Routledge.

Gardner, H. (2011). *Frames of mind: The theory of multiple intelligences*. Hachette Uk.

Garnelo, M., & Shanahan, M. (2019). *Reconciling deep learning with symbolic artificial intelligence: representing objects and relations*. Current Opinion in Behavioral Sciences, 29, 17-23.

Gibson-Graham, J.K. (2002) '*Beyond global vs. local: economic politics outside the binary frame'*, in A. Herod and M.W. Wright (eds) Geographies of Power: Placing Scale. Oxford: Blackwell, 25–60.

Graham, M., Hjorth, I., & Lehdonvirta, V. (2017). *Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods*. Transfer: European Review of Labour and Research, 23(2), 135-162.

Grant, A. W., Seruwagi, L. Z., & Dodd, M. S. (2011). *Artificial Intelligence Through the Eyes of the Public.*

Gray, M. L., & Suri, S. (2019). *Ghost work: how to stop Silicon Valley from building a new global underclass.* Eamon Dolan Books.

Grosan, C., & Abraham, A. (2011). *Rule-based expert systems*. In Intelligent Systems 149-185. Springer, Berlin, Heidelberg.

Guittard, C., Schenk, E., & Burger-Helmchen, T. (2015). *Crowdsourcing and the Evolution of a Business Ecosystem*. In Advances in crowdsourcing, 49-62. Springer, Cham.

Hara, K., Adams, A., Milland, K., Savage, S., Callison-Burch, C., & Bigham, J. P. (2018). *A data-driven analysis of workers' earnings on Amazon Mechanical Turk.* In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 1-14.

Harvey, P., Jensen, C. B., & Morita, A. (2016*). Introduction: Infrastructural Complications. In Infrastructures and Social Complexity*, 19-40. Routledge.

Hatton, E. (2017). *Mechanisms of invisibility: rethinking the concept of invisible work*. Work, Employment and Society, 31(2), 336-351.

Huws, U. (2014). *Labor in the global digital economy*: The cybertariat comes of age. NYU Press.

Irani, L. (2016). *The hidden faces of automation*. XRDS: Crossroads, The ACM Magazine for Students, 23(2), 34-37.

Jackson, S. J., Edwards, P. N., Bowker, G. C., & Knobel, C. P. (2007). *Understanding infrastructure: History, heuristics and cyberinfrastructure policy*. First Monday, 12(6).

Juniper Research (2020). *Voice Assistant Market: Player Strategies, Monetisation & Market Size 2020-2024*, Research.

Karaca, K. (2019). *Basics of Machine-learning and Algorithmic Opacity*. From a lecture in the course "Transformation of Knowledge in a Digital Age" at the University of Twente, The Netherlands.

Kenney, M., and J. Zysman. (2016). *The Rise of the Platform Economy*. Issues in Science and Technology 32 (3): 61–9.

Kurzweil, R. (1990). *The Age of Intelligent Machines*. MIT Press.

Larkin, B. (2013). *The politics and poetics of infrastructure*. Annual review of anthropology, 42, 327-343.

Law, E., & Ahn, L. V. (2011). *Human computation. Synthesis lectures on artificial intelligence and machine-learning*, 5(3), 1-121.

Lease, M. (2011). *On quality control and machine-learning in crowdsourcing*. In Workshops at the Twenty-Fifth AAAI Conference on Artificial Intelligence.

Leonelli, S. (2016). *Data-centric biology: A philosophical study*. University of Chicago Press.

Mateescu, A., & Elish, M. C. (2019). *AI in context: The labor of integrating new technologies*. Data & Society Report.

McCauley, D. (2016). *Artificial intelligence in the real world: The business case takes shape*, The Economist Intelligence Unit Report.

Monett, D., & Lewis, C. W. (2017). *Getting clarity by defining Artificial Intelligence – A survey*. In 3rd Conference on" Philosophy and Theory of Artificial Intelligence, 212-214. Springer, Cham.

Nasteski, V. (2017). *An overview of the supervised machine-learning methods*. HORIZONS.B. Vol.4, 51-62.

Pasveer, B., & Valkenburg, G. (2018). *WTMC Summer School Infrastructure 2018. (WTMC Series on Teaching and Learning; Vol. 2018, No. 02*). Netherlands Graduate Research School of Science, Technology and Modern Culture.

Prendky, J. (2019) *Active Learning. Why smart labelling is the future of data annotation*. Figure Eight Presentation.

Poutanen, S., Kovalainen, A., & Rouvinen, P. (Eds.). (2019). *Digital Work and the Platform Economy*: Understanding Tasks, Skills and Capabilities in the New Era. Routledge.

Ratner, A., Bach, S. H., Ehrenberg, H., Fries, J., Wu, S., & Ré, C. (2020). *Snorkel: Rapid training data creation with weak supervision.* The VLDB Journal, 29(2), 709-730.

Ribes, D., & Finholt, T. A. (2009). *The long now of infrastructure: Articulating tensions in development.*

Roberts, S. T. (2019). *Behind the screen: Content moderation in the shadows of social media*. Yale University Press.

Roh, Y., Heo, G., & Whang, S. E. (2019). *A survey on data collection for machine-learning: a big data-ai integration perspective*. IEEE Transactions on Knowledge and Data Engineering.

Sandwich, W., KRC Research (2016). A*I-Ready or not: artificial intelligence here we come! What consumers think & what marketers need to know*.

Scholz, T. (2012). *Digital labor: The Internet as playground and factory*. Routledge.

Star, S. L. (1999). *The ethnography of infrastructure*. American behavioral scientist, 43(3), 377-391.

Star, S. L., & Ruhleder, K. (1996). St*eps toward an ecology of infrastructure: Design and access for large information spaces*. Information systems research, 7(1), 111-134.

Tubaro, P., & Casilli, A. A. (2019). *Micro-work, artificial intelligence and the automotive industry*. Journal of Industrial and Business Economics, 1-13.

Tubaro, P., Casilli, A. A., & Coville, M. (2020). *The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence*. Big Data & Society, 7(1).

Wang, P. (2008). *What Do You Mean by "AI"?* In Artificial General Intelligence, Vol. 171, 362-373.

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., & Schwartz, O. (2018). *AI Now report 2018*. AI Now Institute at New York University.

Willshaw, D. (1994). *Non-symbolic approaches to artificial intelligence and the mind.* Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences, 349(1689), 87-102.

Winston, P. H. (1992). *Artificial Intelligence (Third edition).* Addison-Wesley.

Cited Websites

Amazon (2020). Help & Customer Service. Retrieved May 12, 2020 from https://www.amazon.com/gp/help/customer/display.html?nodeId=201602230.

Amazon (2020). Amazon Mechanical Turk Requester User Interface Guide. Retrieved June 4, 2020 from https://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/Introduction.html.

Anolytics (2020). Retrieved May 15, 2020 from www.anolytics.ai.

Crawford, K., & Joler, V. (2018). Anatomy of an AI System. Retrieved August 11, 2020 from https://anatomyof.ai/.

Engadget (2018). Data from wearables helped teach an AI to spot signs of diabetes. Retrieved July 20, 2020 from https://www.engadget.com/2018-02-07-deepheart-diabetes-cardiogram-ai.html.

Futurism (2017). Google's Artificial Intelligence Built an AI That Outperforms Any Made by Humans. Retrieved July 20, 2020 from https://futurism.com/google-artificial-intelligence-built-ai.

Harari, Y. N. (2017). The rise of the useless class. Retrieved August 11, 2020 from https://ideas.ted.com/the-rise-of-the-useless-class/.

Infolks Group (2020). Retrieved May 24, 2020 from https://infolks.info.

Russell, A., & Vinsel, L. (2016). Hail the maintainers: Capitalism excels at innovation but is failing at maintenance, and for most lives it is maintenance that matters more. Retrieved April 15, 2020 from https://aeon.co/essays/innovation-is-overvalued-maintenance-often-matters-more.

Samasource (2020). Retrieved June 10, 2020 from https://www.samasource.com/our-team.

ScaleAI (2020). Retrieved May 15, 2020 from https://scale.com.

Taylor, A. (2018). The automation charade. Logic Magazine. Retrieved June 20, 2020 from https://logicmag.io/failure/the-automation-charade/.