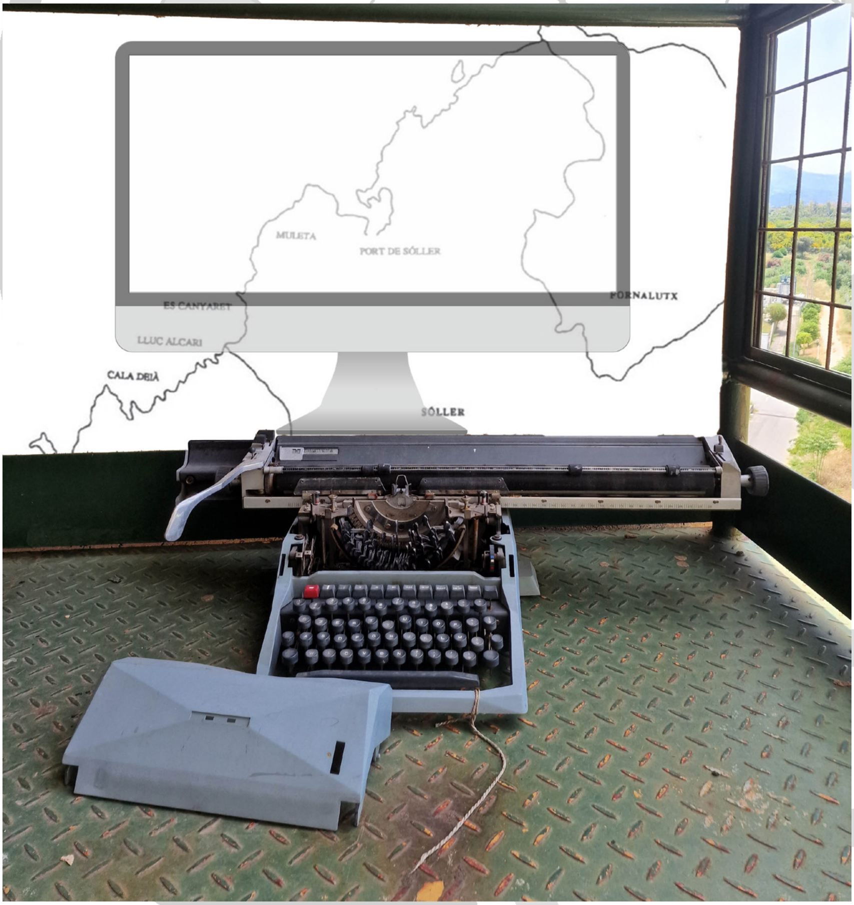
INFORMATION TECHNOLOGIES AND SOCIAL MEDIA: NEW SCIENTIFIC METHODS FOR THE ANTHROPOCENE

Gaetano Sabato, Joan Rosselló (Editors)



Preface by Javier Martín-Vide

IL Sileno Edizioni

OPEN ACCESS
PEER-REVIEWED SERIES
Geographies of the Anthropocene

Information Technologies and Social Media: New Scientific Methods for the Anthropocene

Gaetano Sabato, Joan Rosselló

Editors

OPEN ACCESS

PEER-REVIEWED SERIES

Geographies of the Anthropocene

IL Sileno Edizioni

Information Technologies and Social Media: New Scientific Methods for the Anthropocene
Gaetano Sabato, Joan Rosselló (Eds.)

is a collective volume of the Open Access and peer-reviewed series
"Geographies of the Anthropocene"
(Il Sileno Edizioni), ISSN 2611-3171.

www.ilsileno.it/geographiesoftheanthropocene

Cover: The photo is by Gaetano Sabato. The hand-draw map of the conflict spaces in the coast of Mallorca is by Joan Rosselló-Geli (1995). The graphic project is by Ambra Benvenuto.

International Scientific Publisher "Il Sileno", VAT 03716380781
Via Piave, 3/A, 87035 - Lago (CS), Italy, e-mail: ilsilenoedizioni@gmail.com

ISBN 979-12-80064-36-3

Vol. 5, No. 1 (May 2022)

Geographies of the Anthropocene

Open Access and Peer-Reviewed series

Website: www.ilsileno.it/geographiesoftheanthropocene;

The book series "Geographies of the Anthropocene" edited by the International Scientific Publisher "Il Sileno" (Il Sileno Edizioni) will discuss the new processes of the Anthropocene epoch through the various worldviews of geoscientists and humanists, intersecting disciplines of Geosciences, Geography, Geoethics, Philosophy, Socio-Anthropology, Sociology of Environment and Territory, Psychology, Economics, Environmental Humanities and cognate disciplines.

Geoethics focuses on how scientists (natural and social), arts and humanities scholars working in tandem can become more aware of their ethical responsibilities to guide society on matters related to public safety in the face of natural hazards, sustainable use of resources, climate change and protection of the environment. Furthermore, the integrated and multiple perspectives of the Environmental Humanities, can help to more fully understand the cultures of, and the cultures which frame the Anthropocene. Indeed, the focus of Geoethics and Environmental Humanities research, that is, the analysis of the way humans think and act for the purpose of advising and suggesting appropriate behaviors where human activities interact with the geosphere, is dialectically linked to the complex concept of Anthropocene.

The book series "Geographies of the Anthropocene" publishes online volumes, both collective volumes and monographs, which are set in the perspective of providing reflections, work materials and experimentation in the fields of research and education about the new geographies of the Anthropocene.

"Geographies of the Anthropocene" encourages proposals that address one or more themes, including case studies, but welcome all volumes related to the interdisciplinary context of the Anthropocene. Published volumes are subject to a review process (**double blind peer review**) to ensure their scientific rigor.

The volume proposals can be presented in English, Italian, French or Spanish.

The choice of digital Open Access format is coherent with the flexible structure of the series, in order to facilitate the direct accessibility and usability by both authors and readers.

CONTENTS

8. Artificial Intelligence and Anthropocene

Francesco Mele[1], *Antonio Sorgente*[2], *Paolo Vanacore*[3]

Abstract

"Being acrobats of time" – imagining how the world we live in, the natural environment, art, culture or our scientific knowledge will change, by the actions we are currently carrying out - is a very complex process to describe. In this work we choose a subset of possible actions and we will try to analyze the impact of Artificial Intelligence (AI) on this transformation process. AI is a discipline that perhaps more than any other provides a simultaneous and abundant contribution on two axes: the functional one, due to the number of innovative and original systems it produces at the service of society, and the theoretical and methodological one that has an impact on many disciplinary areas. In other words, we believe that AI is tangibly transforming our daily life, and society in general, but at the same time it is changing the face of many scientific and humanistic disciplines in their theories.

In the chapter we will talk about how AI has improved the methodological apparatus of the human, social and natural sciences, such as linguistics, cultural heritage, medicine and jurisprudence. We will also provide some examples of systems developed with AI methodologies on some application domains. We will cover some specific themes that highlight the critical issues of AI in the anthropocene era. In particular, we examine the problem of technological unpredictability and that of the unexpected results of AI systems. For the latter, we also discuss the problem of regulation, opacity and prediction of the future, which is based on data from the past.

At the end of the chapter we report some regulatory proposals concerning the commercialization of AI systems and, some methodological aspects for the impact analysis of these systems.

---

[1] Istituto di Scienze Applicate e Sistemi Intelligenti "Eduardo Caianiello" of the Consiglio Nazionale delle Ricerche (CNR) – Via Campi Flegrei 34, 80078 Pozzuoli (NA) Italy, Università degli Studi di Napoli "Parthenope" - Via Amm. F. Acton 38, 80133 Napoli, Italy, e-mail: francesco.mele@isasi.cnr.it.

[2] Istituto di Scienze Applicate e Sistemi Intelligenti "Eduardo Caianiello" of the Consiglio Nazionale delle Ricerche (CNR) – Via Campi Flegrei 34, 80078 Pozzuoli (NA) Italy, e-mail: antonio.sorgente@isasi.cnr.it.

[3] Istituto di Scienze Applicate e Sistemi Intelligenti "Eduardo Caianiello" of the Consiglio Nazionale delle Ricerche (CNR) – Via Campi Flegrei 34, 80078 Pozzuoli (NA) Italy, e-mail: paolo.vanacore@isasi.cnr.it.

Keywords: Artificial Intelligence, Anthropocene, Artificial Intelligence Criticalities, Artificial Intelligence Opacity, Artificial Intelligence Regulation.

1. Introduction

In the chapter we will mainly discuss the contribution of AI to the various disciplinary areas concerning the natural and humanities sciences. We will also discuss some critical points of AI in the era of the anthropocene[4] – we wonder if AI defends us from the anthropocene that increases it with its *technological presence*.

AI is a discipline that is receiving a lot of attention at the present time. Eric Sadin in (Sadin, 2019) tries to evaluate the *temperature* related to the interest that AI is receiving in the world. Without hiding his irony, he labels this discipline the "golden calf of our centur" and in a peremptory way he reports that:

> … since 2010, Artificial Intelligence represents the most decisive economic challenge in which to invest with determination and without hesitation. In addition to companies, it is the nations themselves that employ all the means in their power to position themselves at the forefront; this objective has become a major national priority for each of them. First of all the United States, which draws up far-reaching strategic plans, supported in particular by Darpa, the NSA, the Department of Defense, and a myriad of universities and research institutes that benefit from federal grants…
>
> But many nations are no longer willing to come second and manifest a willingness to engage body and soul in this fierce planetary competition. This is the case of China, which aspires to get "on the podium" by 2030, thanks to programs planned in detail, which would lead it to become the undisputed world leader in the following five years.
>
> Canada claims to stand as a "global AI hub" and supports companies and laboratories with the help of generous public

[4] In this chapter we will use the term "anthropocene" often in the sense of a negative path of our humanity that leads to a state of degradation, not only of the physical environment where we live, but of the set of our cultural and social values.

> funds.
> Russia, for decades almost non-existent in the panorama of the electronics industry, plans to become one of the protagonists of this sector, …, Vladimir Putin has in fact declared that "the leading nation in this field will dominate the world" and that therefore "we should avoid leaving the monopoly in the hands of a single nation". The list of countries wishing to try their hand at this promising epic is very long, Israel, Japan, South Korea, … The United Arab Emirates has even gone so far as to set up a ministry for Artificial Intelligence: "Artificial intelligence will be the next big revolution. And we want to be ready"

AI more than any other discipline is characterized:

1. for having provided an interesting increase in innovative artifacts that have been produced using AI technologies (functional contribution of AI)
2. for having inserted new methodological aspects in disciplines that also have robust foundations such as mathematics, engineering and medicine (methodological contribution of AI)

With regard to the first point, AI in recent years has contributed to proposing artifacts with original functionality and a certain usefulness for mankind. If the past millennium was characterized by the arms race by nations, to have prestige and capacity for political control, this millennium was born and grows with the prerogative of possessing the best technology. In particular, at this current moment there is a widespread belief that those who are in possession of technologies such as those of AI will be a country that will have an economic and political competitive advantage over others. Every country right now is drawing up a strategic program for AI, or already has it.

With regard to the second point, namely that of the methodological contribution (which will be discussed in detail in the section 2), AI has contributed to formalizing and conceptualizing knowledge and methods of reasoning of different existing disciplines. In this direction, AI debuted more than 40 years ago with expert decision support systems (especially in medicine) and automatic diagnosis systems. In the last decade, AI has entered with authority in disciplines such as jurisprudence and those related to financial markets. In other disciplines, the methodological contribution of AI has been so decisive as to replace almost entirely the methodological apparatus such as in the domain of Cultural Heritage and in the study of natural languages.

It should be noted that in some disciplinary areas, even without providing a strictly methodological contribution, the inclusion of AI tools has created, and continues to create, a change in procedures and practice in these disciplines. For example, such as the inclusion in some courts of tools to support the decision of the judgment of punishment or acquittal. We believe that this AI contribution should also be considered a methodological contribution. The European community is deeply concerned about the development of AI methodologies. In the section 4 we report the regulatory proposals of the European community for the regulation of the production of products developed using AI technologies.

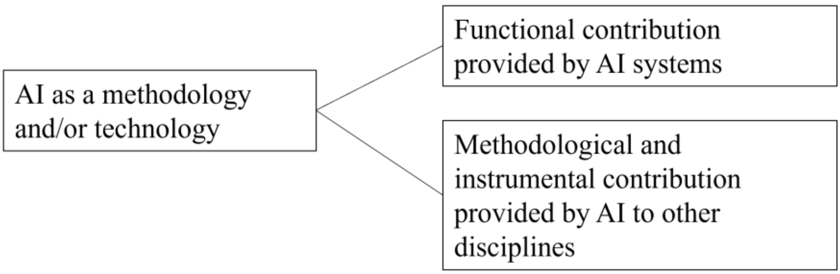


Figure 1 - *Functional and methodological contributions of AI.*

2. Methodological and instrumental contribution of AI to the natural and human sciences

For what follows in this section we will try to maintain, where possible, the distinction between the contribution of AI on the methodological and instrumental level. To better clarify, here are two examples. The introduction of reference schemes or taxonomies in Cultural Heritage (we are talking about the ontologies that we will better specify later) as a classification method is a methodological contribution that AI has provided to the sector. With the advent of the ontologies of AI in Cultural Heritage (see Figure 2), the superintendencies for Cultural Heritage have had to change their way of classifying an object of art. While for example the adoption of AI systems in Law (built with Machine Learning methods) has provided an instrumental contribution in the sector, because it has led to a change in evaluation practice. Although in the same field of Jurisprudence there is to be questioned how much the axiomatic apparatuses of the deontic logics, of an AI nature, have “inspired” the most rigorous formulation of some new law, and therefore have provided a more specifically methodological contribution to the discipline.

In the analysis of the methodological contribution of AI to disciplinary areas we will discuss only some of the main methodological areas of AI: Knowledge Representation, Logic Programming, Natural Language Processing, Machine Learning, Computer Vision and Robotics-AI.

2.1. Knowledge Representation

We open the analysis of the contribution of AI to other disciplines with the Knowledge Representation. This area was the first set of AI methodologies. Since this internal term was coined in the 60s to now, the expression “Knowledge Representation” has gone into disuse for the simple reason that almost all AI methodologies address this problem. The term, however, historically characterized, in a positive sense, the first applications of AI.

The first AI applications that used Knowledge Representation methodologies was called Expert Systems. In those years before the advent of AI many areas of knowledge required initial conceptualizations. The initial goal of this area was precisely to make explicit what was implicit.

Although there was physics that successfully provided models for natural phenomena, there was no methodology capable of making explicit the great knowledge that exists in practice and that leads to problem solving. For example, in the field of medicine there were the skills of doctors to discover the causes of diseases, which required to be codified through systems of rules of thumb. The methodological approach of systems expert in medicine has also migrated to other application areas such as fault diagnosis. Over time, systems that are experts in AI technology have suffered a rapid decline, due to the fact that the knowledge represented in them required great resources to be updated.

Certainly, the significant turning point in the field of knowledge representation took place a long time later, with the advent of a methodology of fascinating perspectives, including theoretical ones: ontologies. Methodology that has established itself in many application areas. Ontologies have played a fundamental role in knowledge management. We can say without exaggeration that there are no social and technological areas where ontologies have not been used.

In order to give a qualitative and non-formal idea of an ontology, we report a simple example in the figure 2, where it can be observed that in this structure there are classes belonging to a taxonomy. Where for each class there is a description that is represented through attributes. In such a structure the individual elements of a domain, represented as instances of the classes, are thus classified. In an ontology there is always a mechanism of prototypical inheritance between a class and its subclasses. This mechanism allows each class to propagate its attributes to each of its subclasses, giving these taxonomies an efficient method of classification.

One of the main tasks to which a nation's superintendency must perform concerns the classification of art objects. The ontologies in the field of Cultural Heritage have provided both an instrumental contribution, but also and mainly methodological. A complex domain that of cultural heritage that involves spatial, temporal and causal knowledge. Each object of art has its own spatial location, its own history made up of events that are connected by causal relationships. AI methodologies could not remain outside the ontologies sector (Bordoni *et al.*, 2013; Bordoni *et al.*, 2016). The methodological contribution of AI has *wiped out* the small and weak classification methodologies existing in this area. The institution of many nations responsible for the classification of Cultural Heritage, first "clinging" to the most famous methodology of the Dublin Core (Kunze & Baker, 2007), then ended up adopting different approaches and formalisms for the ontologies that obviously needed to be integrated[5].

[5] For the problem of the mediation and integration of ontologies, a project SM@RTINFRA-SSHCH (Smart Integrated Infrastructures for Data Social Sciences, Humanities and Cultural Heritage Ecosystem — Italian Ministry of Education, University and Research decree nu. 973 of 25 November 2013) has been activated in Italy. This project was sponsored by the CNR (National Research Council of Italy). This project had an extensive scope of application that did not concern only the field of Cultural Heritage.

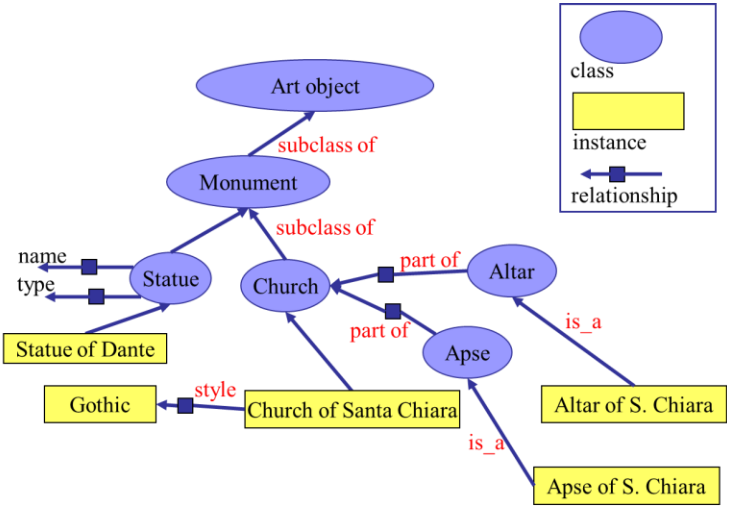


Figure 2 - *Cultural Heritage - a sketch ontology.*

Cultural heritage has taken a large part of the methodological heritage of AI by elevating this sector as a scientifically well-founded discipline. This change has taken place thanks to some new generation operators of the superintendents who have fought against the poverty of method and the arrogance of the generation of operators who preceded it.

Around the 2000s on the axis of AI methodologies of knowledge representation a visionary idea and great perspectives the Semantic Web was proposed (the term is presented in detail in (Tim Berners-Lee & Lassila, 2001)). With this term we mean in essence to provide a semantics to each entity present in the Web, associating a category of belonging, for each of them. This allows web data and knowledge to be easily found and interpreted.

In fact, the application of the Semantic Web (Web, 2021) as a methodology for developing the network (even if not all the web has been brought to this level of operability) has made it possible to facilitate users' access to knowledge and to integrate (semantically) the knowledge of the systems present on the web. The Semantic Interoperability (Interoperability, 2021) and the integration of knowledge sources present on the web has been a significant contribution of AI, where the role of ontologies has been decisive above all in the flexibility and rigor in representing knowledge.

2.2. Logic Programming

Although Logic Programming is on the axis of knowledge representation, its methods make it a unique tool within AI and therefore worthy of being presented separately.

Born around the 60s as the computational form of the concepts of Mathematical Logic. The major proposal in those years was the Prolog (Cohen, 1988) language, from which the Fifth Generation research project was defined, which involved the construction of specialized hardware that reproposed the structures of the computational logical paradigm.

The importance of Logical Programming lies in the fact that it has shown that rigorous axiomatics, concentrated in a few lines of program, are sufficient to model important human activities such as reasoning on actions and events (Kowalski & Sergot, 1986a). In fact, these axiomatics have been used as basic modeling representations of robots moving in ordinary and hostile environments. A separate topic constitutes the representation of spatial relations and related inferences (Stock, 1997). In general, logical programming has been used to model and then simulate almost all human activities that carry out common sense reasoning (Muller, 2015).

But what methodological benefits can a discipline that simulates human behavior give? Certainly, making reasoning explicit and finding out whether or not human rational agents are in possession of specific rules of reasoning, or if some of them apply reasoning containing fallacies and contradictions. Cognitive Sciences have inherited many useful methods from logical programming, especially in the simulation of mental processes.

From the instrumental point of view, Logical Programming has been included in numerous programs also used daily, such as spell checkers, specific expert systems and recommendation systems. And in other application sectors such as relational database management system, expert system, natural language processing, symbolic equation solving, planning and prototyping[6].

2.3. Natural Language Processing Technologies

Natural Language Processing (NLP) is an area of Artificial Intelligence that deals with computational methods and techniques for analyzing and

[6] https://www.easyexamnotes.com/p/applications-of-logic-programming.html

representing texts at one or more levels of linguistic analysis in order to obtain human-like processing for various natural language tasks.

Most of the technologies of NLP have made a contribution (methodological and instrumental) to the study of natural language and the communication.

NLP results have also had a great approval in other sectors, especially in the last decade. There are NLP techniques that have become support tools and/or an integral part of the methods of different disciplines/domains.

The basic techniques of NLP allow the analysis of texts at various levels (Morphological, Lexical, Syntactic, Semantic and Pragmatic) which are the basis of other more complex processes and techniques that describe a class of problems. The main ones are: Text Classification, Term Recognition, Text Summarization, Topic Modeling, keyword extraction, Information Retrieval, Conversational Agents and others. NLP approaches are of great interest to all disciplines that need to analyze texts for the identification of correlations or answers.

Many applications have made it possible to define new research and intervention protocols, or at least alternative protocols for medicine. Just think of the chatbots designed for the triage and initial diagnosis phase, such as Sensely[7], or even the health care chatbots like Amanda Care[8] that monitor patients in order to improve adherence to treatment. In addition, there are systems such as Babylon Health[9] that support the doctor in remote consultations by providing suggestions analyzing the patient's responses, or that suggest ways of life to prevent disease or not worsen the state of a disease like AIDA[10]. These tools are part of Sustainable Development Goal (SDG) 3 – Health and Wellness.

Another area that has benefited of NLP tools is Education. In fact, these techniques are adopted, and support teachers, to improve students' reading and writing skills. In addition to the classic systems for automatic correction or in-depth suggestion for the content different tools have been defined. One example is Cognii[11] , a virtual assistant that engages students in personalized tutoring conversations, providing instant scoring and feedback on written answers to open-ended questions. In the Netherlands, the De-Enigma (Riva & Riva, 2020) project has created a robot with multimodal interaction (facial, body, vocal and verbal signals) for the recognition of emotions and expression

[7] https://www.sensely.com/
[8] https://amanda-care.com/
[9] https://www.babylonhealth.com/
[10] https://www.aidachatbot.it/
[11] https://www.cognii.com/

to teach autistic children of school age. Also interesting is the StorySign[12] project created by Huawey, an application that aims to help deaf children improve their reading skills by translating the text of selected books into sign language. These are some projects of how NLP supports SDG 4 – Quality Education.

On the other hand, from a social point of view, part of the research is currently focused on the predictive analysis of text-based signals, such as those coming from social networks like Twitter or Facebook. Today, textual data is a very rich source of information, and it is growing day by day. Today, most individuals use these platforms to make decisions about purchasing goods, travel, or expressing opinions on social, political and other topics. These predictive techniques are used in many areas from the prediction of the elections of a political candidate to the study of opinions on major issues such as homophobia, racism, bullying, etc., used for the objectives of SDG 16 - Peace, justice and strong institutions.

2.4. Technologies Machine Learning

In 1959 Arthur Lee Samuel, a pioneer of the Machine Learning (ML), reports some studies that "*have been concerned with the programming of a digital computer to behave in a way which, if done by human beings or animals, would be described as involving the process of learning*" (Samuel, 1959).

Learning techniques can be classified into three main categories: *supervised learning*; *unsupervised learning*; *reinforcement learning*.

The *supervised learning* algorithms build a predictive model starting from a set of tagged data (said *samples*). The labels, associated with each example, represent the expected results that the system must "learn" to provide. The labels can have discrete or continuous values; the ML algorithms are called *Classification* algorithms in the first case and *Regression* algorithms in the second one. In this type of learning the training process ("ability to learn") is reported in Figure 3. The sample dataset, called *Dataset*, is splitted into three subsets: the *Training set*, the *Validation set* and the *Test set*. The samples consist of the data for which the system must provide a prediction, and the expected results. The model is trained on data from the Training set and the performance is evaluated using the Validation set. If the achieved performances do not meet expectations, the model parameters are changed (process known as *parameters tuning*), and the training process is repeated.

[12] https://consumer.huawei.com/uk/campaign/storysign/

At a satisfactory level of performance reached, the model is tested on a set of data that did not take part at the training process, named the Test set. As result of the described training, validation and test process a *model* which provides a prediction, in probabilistic terms, from input data (*observations*), is obtained.

In the *unsupervised* learning, the data has no labels. These algorithms try to extract information from the data without an expected result being known a priori. Some examples are the *Clustering*, in which you try to select and group data into "clusters", based on a measure of "similarity"/"homogeneity" between the data; the *Association*, which attempts to identify relationships between data; the *Dimensionality Reduction*, in which the system tries to reduce the data dimensionality, identifying potentials correlations between them.

In the *reinforcement learning* the system does not learn from a Dataset but, acting to achieve a goal, modifies its own behavior (future actions) in function of the feedbacks (*rewards* or *punishments*) provided by the environment with which it interacts.

The instrumental contributions of ML are many and transversal to various sectors including e.g. the Financial Markets and the Jurisprudence.

In the financial sector, and more specifically in the stock markets and the stock market indexes, ML applications are used to predict price trends. In "*Machine Learning for Quantitative Finance Applications: A Survey*" (Rundo et al., 2019) a description and comparison of different ML systems, for quantitative finance with application implications in trading systems and for financial portfolio managements, are reported. The analyzed models are based on both *technical and fundamental analysis* approaches and have the goal of predicting time series by maximizing the accuracy. Technical analysis is based on the assumption that market movements are cyclical (that is, they have patterns repeated over time) and are trained on Datasets containing historical market data. On the other hand, the fundamental analysis seeks to identify the factors that determine market trends, in order to exploit any correlations to predict future market movements. If on the one hand the technical analysis underestimates the variability of the markets, on the other hand, the fundamental analysis can be computationally complex, and this affects the decision rate of the system.

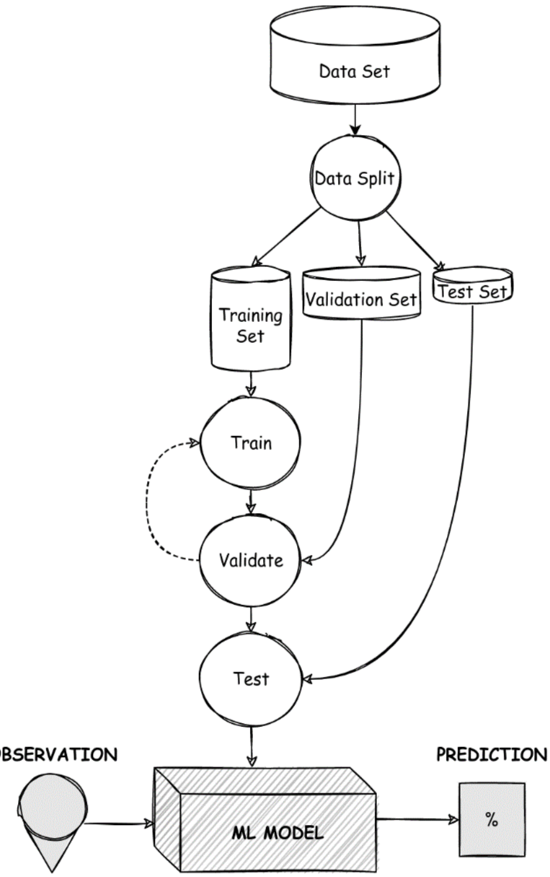


Figure 3 - *Supervised training process.*

Also, in the area of jurisprudence ML-based applications have been developed in support of judges and in judicial proceedings. In 2019, for example, in China, an application called System 206[13] was used in a courtroom to allow a judge and lawyers of the parties to request and quickly obtain documents, expert opinions and videos. In addition to useful in the trial stages assistance systems, such as System 206, decision-making systems have been developed. An article published by the "South China Morning Post", on December 26, 2021[14], a system based on ML capable of making accusations is reported. Shi Yong and his team of researchers trained a model on a dataset of over 17.000 court cases from 2015-2020. The implemented system, which uses natural language descriptions of a case, is able to identify a hypothesis

[13] https://www.chinadaily.com.cn/a/201901/24/WS5c4959f9a3106c65c34e64ea.html
[14] https://www.scmp.com/news/china/science/article/3160997/chinese-scientists-develop-ai-prosecutor-can-press-its-own

and formulate an accusation. The dataset and the system predictions cover the most frequent crimes affecting the city of Shanghai, such as credit card fraud, theft, gambling, dangerous driving, injury, and so on. The research team claimed an over 97% system accuracy.

As is evident in the case of jurisprudence, the instrumental contribution of ML influences, in the practice of use, the methodological approach.

2.5. Computer Vision

Computer Vision is a field of artificial intelligence that deals with methods and techniques for analyzing images and videos in order to allow computers to reproduce human visual functions and processes. Machines can accurately identify and classify entities and then react to what they *see*. In many areas, Computer Vision competes and surpasses human vision. Some Computer Vision techniques involve image segmentation, object detection, facial recognition, image classification, background detection and others.

This AI discipline has also entered various sectors, for example Medicine, redefining some procedures. Many medical diagnoses are based on the study of images and many tools have been developed in the Computer Vision field to help doctors identify pathologies and/or anomalies. An example is X RAIS[15], a platform for medical image analysis developed by Laife Reply that uses 106 different diagnostic methods to support the doctor by automatically suggesting suspicious areas and performing the related classifications. The goal is to reduce the number of misdiagnosis and improve the efficiency of the entire diagnostic process. Midis Ayni Lab[16], is a Peruvian medical project that allows low-cost diagnosis to detect anemia by analyzing images of the eye.

Another contribution of Computer Vision, of great international significance, was made to support investigations to search for missing children. The non-profit organization "International Center for Missing and Exploited Children" (IC-MEC) has announced the launch of GMCNgine[17], a system that helps find missing or abducted children by comparing their photos with those of children online.

There are also human science disciplines that have had great benefits of using Vision Computer technology such as archaeology. With these techniques, systems have been built to support archaeologists in their

---

[15] https://www.reply.com/en/industries/public-sector-and-healthcare/x-rais
[16] https://fairlac.iadb.org/en/midis
[17] https://gmcngine.globalmissingkids.org/

evaluation and classification of the finds found, thus introducing a faster process in the classification phase (Resler *et al*., 2021).

2.6. Robotics-AI

In recent years, AI has provided contributions regarding systems and processes in automation and robotics, obtaining important results such as, for example, in the Medicine sector and in the Automotive Industry.

Medical Robotics includes applications to support various activities, from diagnosis and prevention, to surgical operations, physiotherapy and rehabilitation practices. Recent technological development has made possible to create micro-robots capable of exploring the human body, with ever greater precision, in order to diagnose and prevent diseases. Even in operating theaters, through robotic arms and camera systems controlled by surgeons, sometimes even remotely, robotics makes an important contribution to the implementation of precision interventions. In rehabilitation and physiotherapy practices, in cases of permanent or temporary disabilities as a result of trauma or disabling pathologies, exoskeletons and, more generally, wearable devices have been created to allow correct mobility and/or the recovery of compromised functionality of limbs and hands.

Among the various objectives of AI applied to the Automotive Industry are the optimization of the use of transport infrastructures, the improvement of mobility, the support for people with disabilities, the minimization of risks through active safety systems, the transport times and, consequently, the energy consumptions. The first attempts to build driverless vehicles can be traced back to 1925, when the US Army electrical engineer Francis P. Houdina modified a Chandler automobile by equipping it with a radio antenna and electric motors for remote control its movements via radio control. Over the years, several prototypes have followed one another, with increasing levels of automation, with the aim of achieving fully automated driving control without the need for any human intervention. Among the first examples of vehicles capable of automatically processing signals from the environment, through cameras and sensors, is the German prototype "VaMoRs", built by the engineer Ernst Dickmanns of the University of Munich in 1985. In 1994, the "VaMP" and the "Vita-2", also made by Ernst Dickmanns in collaboration with Mercedes Benz, were self-driving cars based on computer vision techniques and were tested for over 1000 km, with an average human intervention required estimated at one every 9 km. In 1998 the ARGO project, led by Prof. Alberto Broggi of the Department of Information Engineering of the University of Parma, was developed as part of the Transport Project 2 of the Italian National Research Council[18]. The project saw the modification of a Lancia Thema, which was equipped with two video cameras for processing data from the external environment. The vehicle traveled about 2000 km in six days, and an autonomous driving time was estimated to be 94%. Subsequently, numerous prototypes were made up to commercial vehicles, both automobiles, such as the popular products of Tesla Motors, as well as "heavy" vehicles, for the construction and public transport.

3. Criticalities of AI in the Anthropocene Era

In the previous paragraphs we have discussed the transformations due to the methodological and instrumental contribution of AI in other disciplines. This benefit of increasing method and theoretical improvement is certainly positive in the path of growth of knowledge of humanity. It is clear that an extension of method can lead to the development of new technologies and the development of new systems. But, it is on the latter and their functionalities that careful control must be carried out, not on the development of disciplines. They raise concerns, therefore, about the impact that determine the functionality of new AI systems in the anthropocene era, where one wonders what individual and social actions such systems can cause. Are there any obscure changes that our society undergoes (or could undergo), with (or without) awareness, for the injection of new features labeled "intelligent" or that independently perform some tasks that previously only man performed? We have grouped in some themes (not exhaustive), the answer to this question.

3.1. The application unpredictability of AI systems

In this paragraph we will discuss the unpredictability of the application directions that AI technologies can take. The latter problem is different from the problem of "unexpected results" of AI system that we will discuss in the next paragraph 3.2.

In (Tamburrini, 2020) the author seems to agree with Ellul's thought that the direction of technology is really unpredictable. The work gives the example of IBM's Watson system, which is very relevant to the ongoing discussion. Watson was originally designed to participate in Jeopardy!, a

[18] Consiglio Nazionale delle Ricerche (CNR)

general-purpose television quiz show, where he scored a victory over human competitors. The system was built with AI learning technology using natural language technologies. A domain, that of the game, that did not suggest dangers of use. However, it happened, using the same methodological apparatus and AI approach, was developed on a new application on another domain considered ethically sensitive – such as that of purchasing advice.

We agree that unpredictability is a characterizing factor for every technology and especially for AI, but we also believe that worry should be directed to the type of use of a methodology, without limiting the discovery of new applications that can be generated by the methodology itself and without also placing constraints on the extensions of it. We believe that a methodology in the design phase can also allow a glimpse of some application field that requires a subsequent precautionary analysis of use. But it is on this use that a certain filter and control must be made, not on the development of the methodology.

We report the case of Hacronym (Stock et al., 2002) a research project funded for the study of computational humor. This project had among its potential applications in the educational field, since humor presents itself as a very effective form of communication for teaching. Hacronym presented, however, as many possibilities of being used in the commercial field of advertising and purchasing advice. What to do then? Not undertake a fascinating research challenge that studied man in one of his most creative activities such as humor? The research on computational humor was developed and coordinated by the Bruno Keller Foundation and completed around 2001. This research reported flattering theoretical results in humor modeling (Stock *et al*., 2002) that enriched many different disciplinary areas such as linguistics, logic, cognitive sciences, and of course AI itself. To our knowledge, research related to Hacronym did not lead to the development of e-commerce systems.

We believe that with a good methodology of impact analysis and control such as the one we will present below (section 4.1) can greatly reduce the problem of "unpredictability of AI systems". Indeed, we believe that this aspect, if properly managed, can become a positive feature for the search for original application opportunities. We fully agree with sentence reported in (Cucchiara, 2021): "the way is to regulate not research on AI, but its applications and final products".

3.2. Unexpected executions of AI systems

Unexpected executions of AI systems[19] does not concern the application unpredictability problem, the latter described in previous paragraph. The latter is inherent in the direction that a research can take and therefore not foreseeing in which sector such a methodology will be applied. An unexpected result, however, concerns the fact that a system that has been defined to operate a certain way, against the intentions of its builders, behaves differently. A popular chatbot, built with AI technology called Tay, was designed to follow language patterns in order to be friendly and reasonable. In some of his executions, however, he displayed racist and sexist behavior[20]. The problem of unexpected results is not an easily solved problem, especially due to the fact that AI systems have achieved a high quality of performance of the tasks they are required to fulfill (especially those concerning aspects of perception, reasoning and learning - typical of AI) and, what happens is the higher the quality of the tasks, the more difficult it is to discover unexpected behaviors of AI systems.

In this context we believe that in the verification of the functioning it could be of some benefit to design and develop systems that have explanatory modules (see forward the paragraph 3.4). The explanations make it possible to analyze the inferential paths of a certain system and therefore allow the correction any unwanted or even erroneous behaviors.

For the problem dealt with in this paragraph we believe that it is also useful to carry out adequate design methodologies[21].

---

[19] The famous American statesman Henry Kissinger written an article "How the Enlightenment Ends. Philosophically, intellectually — in every way — human society is unprepared for the rise of artificial intelligence" (Kissinger, 2018). The article contains three main points of discussion:

> First, that AI may achieve unintended results. Second, that in achieving intended goals, AI may change human thought processes and human values. Third, that AI may reach intended goals, but be unable to explain the rationale for its conclusions.

The first point is the argument of discussion of this paragraph, while the third point will be discussed in paragraph 3.4.

[20] https://dailywireless.org/internet/what-happened-to-microsoft-tay-ai-chatbot/

[21] For example, the *i** (Yu, 1997) framework was developed for modeling and reasoning about organizational environments and their information systems and can be adopted in the early-requirements modeling step.

3.3. Difficulty for regulating AI systems

There are fields of application of AI methodologies, which are crucial for addressing a regulatory discussion of the use of such systems. There are areas that require urgency and particular attention and are those related to the protection of life and human health in general - specifically, we believe that priority should be given to the arms, medicine and transport sectors, where various technologies have been used robotics-AI (see section 0).

With regard to the regulation of AI systems, the points reported in (Cucchiara, 2021) are expressed very clearly and are fully shared by us, even if we believe they must be addressed with varying degrees of urgency. The points are: (1) No AI weapons, (2) Identifying accountability, (3) Understanding the nature of Intelligence, (4) Privacy, and (5) Human control over generalization.

The first point that the author labels as a priority is expressed with the slogan: "No AI weapons". We agree, in no uncertain terms: "to ban the weapons based on Artificial Intelligence and to intervene on global level (worldwide)". For this, we believe that a research ban on associated methodologies in this topic would also be necessary. However, we believe it is useless to debate on the topic of "armaments ethics" trying to discriminate between autonomous functions and those of human competence. "Ethics" and "weapons" cannot exist in the same context of discussion since they are terms that are mutually exclusive, whatever meaning one may choose for the term "ethics".

The second point contained in (Cucchiara, 2021) is "Identifying the liability for damages from the use of AI systems". It is a very difficult point to solve, because it involves not only a discussion on what AI is and the functionality of its systems but also aspects of a social nature because it is necessary to address the problem of responsibility. For the first discussion we believe that the proposed analyzes on the levels of autonomy both in medicine (Yang *et al*., 2017) and for vehicle control (SAE, 2018) are valid, in order to better understand from this decomposition which are the functions that the machine must perform and which are those of man. This analysis also helps to understand which anomalies are labeled as "machine malfunction errors" or even "design errors". With these analyzes as input, it is possible to switch more clearly to the assignment of responsibilities.

This last question of a behavioral-social order constitutes the heart of the problem. Various responsibilities and actors gather on it. The manufacturing companies together with their programmers, although the need to sell a product, tend not to take any responsibility, while users of AI robotics systems require guarantees of correct operation and do not wish to have responsibility for any adverse situations. Independent groups do not seem to have the right reasons to start initiatives to draft liability regulations. Consider the case of Google's proposal to set up an independent ethics committee, ATEAC, aimed at discussing cutting-edge AI initiatives that quickly failed[22]. The proposal was withdrawn in a few days due to protests relating to the existence of conflicts of interest of some members of the committee.

We believe that interesting initiatives can only arise at the government level, where there would be the right reasons for drafting regulations for the responsibility of the use of AI robotics. It is the government's responsibility to promote a reduction in road accidents, or to include more effective surgical tools in hospitals.

The applicability of the rules of responsibility presents a last not negligible problem, which in this pandemic period has emerged in all its facets, that is, the fatigue attitude and sometimes of rejection of people to social regulations. To this last problem some elements of resolutions have been the object of the attention of the European community which has issued a documentation in which it invites all potential users of AI systems to have confidence in this technology, a confidence that for the European community can only be acquired with a better understanding of AI technologies.

We believe that point (3) concerning the "Understanding the nature of intelligence" also falls into this last discussion. We will not address point (4) in this paragraph because we believe it is a general problem and it is not a discussion topic of this work in which we are discussing, in particular, AI systems.

Regarding point (5) Human control over generalization, we believe that the generation of systems that always try to generalize and organize everything into classes, can lead to classification problems, generating problems, in some cases, even of racial and ethnic ones. This can arise because in trying to improve these systems more and more features are considered and they can bring into the model some biases present in the training data (as also reported in the explanation section) and that, even if correct, may not be always valid in a society that is constantly evolving. An example is *predictive policing*, a system that predicted the likelihood of crime in certain areas and thus focus police attention. After a first phase it was abandoned as a system because it was discriminatory, the areas identified were mostly areas where Latinos or African Americans reside. According to (Cucchiara, 2021) generalizations, Aristotle taught, make sense if there are

[22] https://www.bbc.com/news/technology-47825833.

general postulates ("*All men are mortal*"), but, since these are few, we must avoid relying on systems that could to say that all men with pointed mustaches are artists or that beautiful women are stupid. Or much worse. You can have systems that can classify artists, stupid people or dangerous areas but they must provide explanations and that take into account characteristics and properties that are not discriminatory for the person.

3.4. The problem of explaining AI systems

One of the open questions related to AI, in particular with the intensive use of systems based on Deep Learning, but also ML, is the problem of *explanation*. We now have systems with high predictive or classification capabilities, but *opacity* has also grown (Gunning *et al*., 2019). This is because most of the models are of the black-box type, so they that do not allow you to inspect the process and therefore the choices/decisions made by the system are not understandable. This is especially important in areas such as medicine, defense, finance and law, where understanding decisions and building trust in algorithms are critical [*ibid*]. As reported by authors in (Guidotti *et al*., 2018), it is not only a transparency problem, in fact the ML and DL models learn from examples and if these examples provided hide prejudices and defects (bias) then the algorithms will suggest unfair choices, for example discriminatory and racist like Compas-Correctional Offender Management Profiling for Alternative Sanctions (Skeem & Eno Louden, 2007). Compas is a system that has been used to predict the risk of criminal recidivism by some US courts to support judges for release claims. A study carried out by some journalists[23] showed that the system had a strong racist bias. In fact, black people were assigned twice the risk of whites even if they were in the same conditions. This implies that the model has inherited a bias from the data, meaning that the data used for training are *biased* towards black people. This system, even if properly trained, considered unethical characteristics of people and was unable to provide an explanation of the assessment made. Having systems that *explain* their evaluations are able to give clues about the choices made by helping experts, stakeholders to make better decisions.

The problem of explanations does not only impact the quality of the functioning of a certain system but in some cases it is decisive for the adoption of the system itself. A sentence of the Lazio Regional Administrative Court

[23] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

of 2019 (summarized in (Numerico, 2021)) rejects the use to adopt an algorithm for the choice of a decision to transfer some teachers from Puglia to Lombardy. The Lazio Regional Administrative Court accepts the rejection of the transfer provision, also raising questions of principle on automatic evaluation systems in general. The main argument for the acceptance of the appeal reside in the fact that the ruling issued by the program is based on rules that are not made public and cannot be subjected to evaluation, neither on the method, nor in the result obtained. In other words, it does not give the res judicata the possibility to oppose the sentence, nor the judge to issue a judgment on the validity of application of the law itself.

3.5. Epistemological opacity of AI systems

In the various disciplines the formulation of the theoretical apparatus passes through a mechanism of abstraction of concepts. This mechanism often concerns a labelling activity using abstract terms that tend to associate categories, ie references, to relationships and processes existing in a domain.

At the beginning of its path, AI started from the representation of knowledge with the aim of making every type of knowledge explicit. For a long time, AI has played the role of a good archaeologist who unearthed hidden relationships, complementing theories with elements of abstractions that were lacking in certain disciplinary systems. With Expert Systems, AI has made the hard-wired knowledge of doctors explicit. From there began a process of formalization and theorizing of abduction which took its most rigorous form in its logical formal representation.

The elegant and rigid formalizations of classical physics do not take into account concepts belonging to the common sense of everyday life. The modeling of natural phenomena must not only be based on formal notions of position, acceleration and force, but also needs to represent behaviors, which require qualitative notions, often contained in expressions in natural language. In physics, for example, notions such as trajectory, speed, acceleration and force have been theorized, but there are no formalizations of the concept of action or event - these are present instead in explicit axiomatic logic programming - such as the Event Calculus (Kowalski & Sergot, 1986b; Miller & Shanahan, 2002) - which have paved the way for innovative applications in Robotics.

And again, in physics the concept of causation is present only implicitly in the natural systems that it describes, but it has never addressed the notion of mental causation, which is the basis of all cognitive activities of human rational agents.

On the linguistic side, AI has overlapped with Chomshy's transformational theories (Chomsky, 1975), where starting from the aspects of representation of the grammar of a language, he founded a new discipline that of computational linguistics (Hausser & Hausser, 2001; Mitkov, 2004), a discipline that has been highlighted not only for the extensive results reported in the applications, but above all for having created a new theoretical and reference approach to natural languages.

Ultimately an AI that accustomed us to surprise us with speed and rigor entering many different disciplines. Along a path that has been proposed as a sort of new epistemology of science that for every even small AI application, has provided methods (even if partial) to represent theories that are also very different from each other.

In this path of AI has suffered, in our opinion, a stop, at least one of the branches in which AI has continued - that of Big Data and Machine Learning, a sector where while proposing applications with interesting features, it has not continued on the path of integration and extensions of existing disciplines. With the new approaches of AI, systems have been developed *imprisoning* the knowledge in a sort of *black box*, where it is difficult to build a process of explanation of the functioning of a system and, it is not possible to identify rules, abstractions and concepts that are the basic constituents of theories.

At this point we would like to say that the AI approach based on Machine Learning or Big Data produce epistemologically opaque systems.

We will not construct a formal definition of epistemological opacity of the theories generated by AI systems – it is beyond the scope of this work. However, we will show how the notion of opacity given by Humphereys (Humphreys, 2009) can be used for processes, to suggest some criteria for evaluating the epistemological opacity of a theory.

We can assume that a set of programs that use AI methodologies evokes (opaquely or not) a theory. We make no assumption of how this mechanism happens. However, we can reasonably assume that every AI program is in effect a process, just as every process is made up of parts, so is a program.

At this point we will use the notion proposed by Humphereys formulated for the opacity of processes:

> A process is epistemically opaque relative to a cognitive agent *X* at time *t* just in case *X* does not know at *t* all of the epistemically relevant elements of the process. A process is essentially epistemically opaque to *X* if and only if it is impossible, given the nature of *X*, for *X* to know all of the epistemically relevant elements of the process.

We can think that this definition can be applied to AI programs, making a substitution of "AI program" instead of "process". It is true (as assumed for processes) that an AI program is made up of parts. And that there are parts of a program that are epistemically relevant and others that are not. We report two different parts P_1 and P_2 belonging to two different systems (programs) S_1 and S_2, of which P_1 and P_2 evoke the same theory, but where P_1 is not opaque while on the contrary P_2 is. The part of theory evoked is that relating to the rule of transitivity, that is:

RT: For all events E_1, E_2, E_3 where E_1 precedes E_2 and E_1 precedes E_3 then E_1 precedes E_3

In S_1, this rule is represented by the program part P_1 as follows:

$$\forall E_1, E_2, E_3: event(E_1) \wedge event(E_2) \wedge event(E_3) \wedge$$
$$prec(E_1, E_2) \wedge prec(E_1, E_3)$$
$$\Rightarrow prec(E_1, E_3)$$

For the rule 1 part of S_1 theory exists (it probably exists, and later we will explain why), a recognition that is epistemically not opaque. While, for the P_2 part of the theory in S_2, it is epistemically opaque – it is a hypothetical inductive P_2 program that starting from a corpus C_n of n statement of events *brings out* the *RT* rule. Namely:

$$C_n\ P \Rightarrow Tex$$

The example given helps us to understand what are the terms to be taken into account to evaluate if some programs and therefore theories are opaque. An explicit and declarative formulation of the parts of a program, where the use of entities such as universally quantized variables (for every X,Y : $F(X, Y)$) as present in rule 1, plays an essential role for the recognizability of the non-opacity of a part of a theory.

The multitudes of instances and *obscure* processes in ML systems, oppose the process of recognizability of the detection of non-opacity. Such systems therefore do not provide theoretical increases and interrupt a path of increasing the theories to which AI had accustomed us.

The aspect of fear, and of anthropocene due to AI, arises precisely from the coming to power of these immense and "obscure" masses of data that constitute only instances (not knowledge) of real-world relationships and processes - instances in search of a clarifying conceptualization.

In other words, with such systems, AI returns to the point where it started, when it proposed to replace databases with knowledge bases, when we listened to seminars in which the difference between information and knowledge was emphasized.

A reset and then build other epistemological pathways? Or a definitive renunciation that surrenders to the slogan: large databases (the Big Data) work well, no matter if I give up increasing my theories and knowledge about the world?

3.6. AI systems, the risk to confirm the past when they making predictions

Prediction of the future is an operation that is performed in many types of applications in which AI systems operate. We believe there are some of them that present a troubling social action: making predictions that confirm people's past behavior. In general, we ask ourselves how we can move away from the anthropocene direction in which we are going, if people, and everything around us, at this moment confirm the erroneous choices of yesterday? Our poor health now is determined by the choice of our eating yesterday; today's polluted air is determined by yesterday's choice of fuel; our bad rulers today, from the votes made to the last elections; the bad movies or fiction that we watch now, from the fictional products we saw in the past television season. Many application areas of AI, such as recommendation systems, base their operation on recording the behavior of people, physical and social phenomena that occurred in the past. They provide purchase suggestions based on future forecasts of people's liking. However, recommending products can guide our judgment, focusing our choices on a limited set of elements, excluding for us useful or preferred products. It is therefore important to understand the operating mechanisms of these AI systems.

Let's examine the latter in detail, then looking for potential corrective actions. Let's take for example a real system RS where we can associate transformations $M_1 \rightarrow M_{11}, M_2 \rightarrow M_{22}, \dots, M_i \rightarrow M_{ii}$, between the states $M_1, M_{11}, M_2, M_{22}, \dots, M_i, M_{ii}$ where RS can be found. Let's leave out the discussion of what the representation of states and transformations should be: if we have all the tools to do it, if we are good at doing it, if there are limits to such a process. In this discussion it doesn't matter.

We emphasize the fact that $M_x \to M_{xx}$ are all possible transformations, that is, those registered and classified, those that have occurred and not registered, and those that we (or any person) do not know.

We then have a corpus of data where a subset of transformations (which we indicate with SM_j) actually occurred in RS have been classified in past time. Let us also imagine that there is a forecasting system P which from a state M_x at time t provides a forecast M_{xx} at time $t + 1$.

If we think to realize a prediction system P only on the cases actually occurred SM_j (subset of all those potentials) then, in the reality, some transformations can occur that are not predictable by the system P.

An example of predictions on all is present in e-commerce systems where user behaviors are classified at a time t_1 and some purchases are suggested at a future time t_2. The choice does not take into consideration whether the product may be useful or the user may like it at time t_2, but on a classification made on the behavior of the same user, or of a user similar to the latter, at a time t_1.

Basically, systems that predict future events that are based only on facts or behaviors that occurred in a previous time, are to be avoided, because they cannot make predictions about events that did not happen, that is, they do not allow a choice to be made on all the possible predictions.

Theoretically, what has just been stated consists of the gap between scientists and politicians approach in the context of the current pandemic. Scientists, although often do not provide a strictly causal explanation of the pandemic event, declare that it could have been avoided if precautions were taken in the interactions between humans and the animal world, while politicians speak of Covid as a completely unpredictable tsunami. For them, nothing that could be observed at that moment gave a glimpse of what would happen. The former base their beliefs using causal laws, the latter perform a sort of induction based only on confirming what has already happened.

3.7. Anthropocene transformation loop of AI technologies

The question of proposing objectives that lead to functionalities that are "useful" or so-called "advantageous" for humans, society or the environment and, at the same time, constraints on the functionalities or methods of use of AI systems, highlights a very topical problem: whenever an *Sx* system defined with certain functions, methods of use and construction specifications is used to obtain advantages for a specific task, it is necessary to highlight the physical, social and environmental context in which the *Sx* system operates, in order to anticipate and reduce any risks or collateral damage that *Sx* could cause. In other words, for these types of problems it is necessary to identify which subsystem of a given physical or social system *Tx* which benefits from the so-called advantage in using *Sx* and which other entities of the same *Tx* system which on the contrary are disadvantaged.

The problem for some systems, such as AI ones, could become complicated if other control systems (perhaps still defined with AI techniques itself) are needed to reduce any risks.

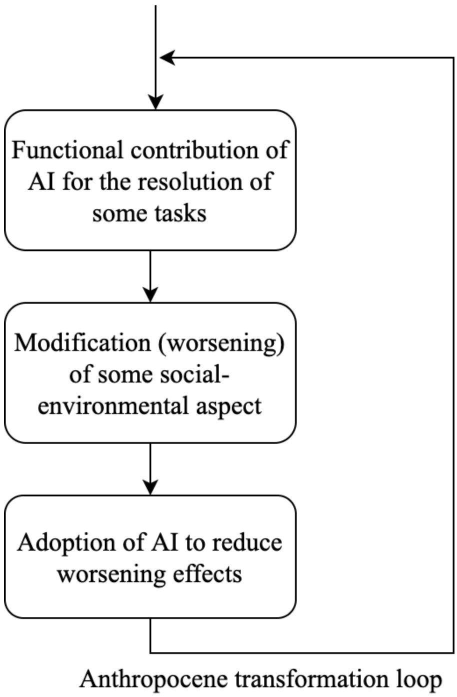


Figure 4 - *Anthropocene transformation loop.*

In an elegant way, in one of his articles, Peyron (Peyron, 2021) questions himself the subject in the following way:

> Artificial Intelligence transforms the society, economic and political relations: if our action is aimed at reducing, limiting, mitigating or reversing the misuse of the human presence in the ecosystem, the so-called anthropocene, the question that arises is

whether to increase the anthropocene is the correct way to reduce the anthropocene[24].

An example of the lats statement is reported in *Energy and Policy Considerations for Deep Learning in NLP* (Strubell *et al*., 2019) where regarding the adoption of NLP systems, in particular with Deep Learning techniques, the impact in terms of *CO_2e* emissions is analyzed[25]. The **Errore. L'origine riferimento non è stata trovata.** shows the estimated cost of the training process, for some commonly NLP models, in terms of emissions *CO_2e* and cost, in euros, by adopting cloud computing platforms.

Table 1 - *Estimated cost of training a NLP model in terms of CO2e emissions and in-cloud compute cost (EUR).*

Model	Power (W)	Hours	$kWh \cdot PUE$	CO_2e (lbs)	In-cloud compute cost (€)
$Transformer_{base}$	1.416	12	27	26	36 - 124
$Transformer_{big}$	1.515	84	201	192	256 - 867
$ELMo$	518	336	275	262	382 - 1.302
$BERT_{base}$	12.042	79	1.507	1.438	3.317 - 11.116
NAS	1.515	274	656	626	833.866 - 2.831.171

The table shows us how the massive use of some AI techniques, which we use daily today, have a negative impact on the environment. So, for the proposed technologies, in particular for AI, it is necessary to leave the anthropocene transformation loop highlighted in the Figure 4). In (Sadin, 2019) the authors provide a suggestion that could lead to the resolution of the problem posed: the cultural profile. The latter consists in choosing a more meaningful use of technology in such a way that the latter pushes the behavior of individuals in a more ecological direction.

In particular, we believe that a solution is the adoption of methodologies for impact analysis, which must become an integral part of AI projects. Regarding this point, we postpone the discussion to the section 4.3.

[24] As one can see, Peyron uses the term anthropocene in the sense similar to that given by us in the footnote of the introduction.

[25] Wikipedia: "*Global warming potential (GWP) is the heat absorbed by any greenhouse gas in the atmosphere, as a multiple of the heat that would be absorbed by the same mass of carbon dioxide (CO2). GWP is 1 for CO2 (. . .) Carbon dioxide equivalent (CO2e or CO2eq or CO2-e) is calculated from GWP. For any gas, it is the mass of CO2 that would warm the earth as much as the mass of that gas*".

4. Remedies and attempts to control AI technologies

With regard to the AI functionalities, the game is played on two tables: that of directives that lead to the creation of systems that exhibit functions (understood as the capacity of artifacts/systems) that provide an advantage for humans and the environment in which it lives, and to provide restrictions on the generated functions.

The Italian AI scientific community took care of the first objective, elaborating an interesting document submitted to the current Italian Prime Minister which also reports the existing AI plans of other countries in the world (document that we present in the paragraph 4.1).

Instead, about the second one, relating to restrictions, the contribution came from the European community with a regulatory document based on the risk analysis for the marketing of AI systems (document that we present in the paragraph 4.2).

In the final paragraph (4.3) of this section we report some basic concepts to build methodologies for the impact analysis of AI systems, during the design, development and verification phases.

4.1 An AI strategy proposal for Italy

The Italian AI scientific community in a document (Semeraro *et al*., 2021) submitted to the Italian political community (Council Presidency) entitled "Artificial Intelligence for Sustainable Development" took entirely as a reference the objectives of the 2030 Agenda for Sustainable Development[26] proposed in September 2015 by more than 150 international leaders in a meeting at the United Nations to contribute to the global development of the planet.

In the Italian document on the massive adoption of AI for sustainable development, an analysis of the impact of AI was carried out on all 17 objectives proposed by the 2030 Agenda. Objectives that we explicitly list: *No Poverty; Zero Hunger; Good Health and Well-being; Quality Education; Gender Equality; Clean Water and Sanitation; Affordable and Clean Energy; Decent Work and Economic Growth; Industry, Innovation and Infrastructure; Reduced Inequality; Sustainable Cities and Communities; Responsible Consumption and Production; Climate Action; Life Below*

[26] https://sdgs.un.org/goals

Water; Life On Land; Peace, Justice, and Strong Institutions; Partnerships for the Goals.

Each goal is enriched by the contents of two key attributes: "what AI can do to achieve the goal" and "what are the dangers and risks to avoid" .

4.2. Guide Lines for commercial AI products

The European community has provided guidance based on risk analysis, proposing to group commercial AI products into 4 risk levels (Commission, 2021):

- a first level consists of AI systems that use subliminal techniques to distort a person's behavior, causing physical or psychological harm to that person or to others;
- a second level labeled as high risk is defined by the functioning of certain AI systems that can have negative repercussions on people's safety or their fundamental rights;
- a third level known as limited risk includes those systems that must be subject to minimum and precise transparency constraints, such as chatbots and voice assistants. These systems must ensure operation such that those who interact must be able to be aware that they are interacting with a machine, in order to be able to decide with full knowledge of the facts whether or not to continue using them;
- a fourth level in which the risk is considered minimal, for the safety, rights and freedoms of citizens. This last category would include maintenance systems, spam filters, and video games developed with AI techniques.

In this regard, always in the Ethical Guidelines on AI (Group, 2019) of the European Commission, it manifests the vision in which, to be ethically correct, the AI must be reliable, compliant with the laws and must comply with the following 7 conditions:

1. human supervision of AI systems, to ensure respect for fundamental rights and the well-being of the user;
2. robustness and safety, such as the safety and reliability of the algorithms and the degree of effectiveness and efficiency of the control systems in the event of hypothetical illegal operations;
3. privacy, control and data management;

4. transparency to guarantee the traceability of the systems and to demonstrate the operations carried out by the algorithm;

5. diversity, fairness, absence of discrimination: artificial intelligence systems should take into account the different and distinct human skills and abilities, at the same time guaranteeing free access to these tools to everyone;

6. social and environmental well-being, that is, always paying attention to the impact on the environment and the social order, promoting the use of AI only where its use can guarantee sustainable development;

7. responsibility, i.e. continuous systems verification, both internally and externally.

4.3. Methodologies for impact analysis of AI systems

The loop presented in 3.7, the directional unpredictability of AI technologies and, the problem of unexpected behavior of systems can be a cause for concern if you think that there are no remedies to make reductions in "quantities of anthropocene" that can be inserted at some point in the development of AI systems.

In addition, a loop reported in 3.7 gives rise to the suspicion that there may be many other loops of a non-technological nature whose existence we do not know. Sadin in (Sadin, 2019) warns us that industrial production is no longer sensitive to the need to carry out multiple and meticulous quality tests. Indeed, it declares that at the moment the trend is that there is:

> almost no discrepancy between design and marketing. Competitive pressure and the primacy of immediate return on investment prevent the slightest latency period, as well as any concerted assessment of the value and relevance of products. Research and development units must prove without delay and relentlessly that they are levers of profit.

However, this does not always happen. AI systems enter our daily lives not only through a strong bet determined by market reasons of unscrupulous companies. But also, through national and territorial research programs funded by government projects. It is these types of programs that we take as a model for controlling the divergence of any technological loops. In recent years, various AI technology projects have been funded, where impact analyses have been associated as a mandatory requirement of the project. A real activity that takes place in parallel with the development of the project itself that goes through the various phases of design, development and analysis of the final result. An example of this type of project is the one that was launched in Italy in the year 2020-2021 by the San Paolo Foundation call, the "Artificial Intelligence, Art and Culture", where the candidate projects were formed by mixed academic, research, business and industry groups. And where, as a design requirement, there was that of in the activities an impact analysis methodology to be applied in all phases of the project. We believe that these latter methodologies are perhaps the only way to make rigorous assumptions of social change.

We will not report what the specific impact methodologies are, but we believe that this chapter is the right context to provide a brief presentation of the basic components. We will exemplify this presentation taking as a reference the AI projects that are composed of mixed research and business groups, which have as their purpose the development of AI products with not only profit-making purposes, but which must by choice or by project constraint perform a social impact assessment.

They play an important role in the construction of an impact chain the set of stakeholders, i.e. all types of natural persons or social entities involved in the development or use of an AI system that is intended to be designed and developed.

The choice of the sets of stakeholders referring to the development project is decisive for identifying the indicators through which the impact assessment will be built and formulated, which we repeat must have a quantitative evaluation characteristic.

In order to formulate an impact value chain, it is first necessary to define the concepts of input, activity, output, outcome and impact (Etica & Consulting, 2016), on which the selection phase of the indicators will also depend. For the interest to AI system projects, we believe that two phases require specific attention:

1. outcome (results) - are all the changes, positive and negative, both short-term and long-term, that occur on the lives of the recipients of the realized AI system. The outcomes are therefore the benefits obtained and also the negative effects verified as a result of the adoption of the system. The outcomes can be short or long term depending on the social needs to be met and the functionality or service provided by the system. They can be direct (reasonably direct consequence of the system or service on the life of users) or indirect (indirect effect on the life of the beneficiary or other natural persons or body of the company). In addition, outcomes can be expected or unexpected, i.e. results not expected after the adoption of a certain system.

2. impact - is the part of the outcome related to change. The impact is therefore a measure of outcome net of the essential changes, which would have occurred equally even without the use of the AI system realized. The impact measurement therefore represents the actual ability of the use of a certain AI system to cause expected changes.

In these impact assessments the key elements to be considered are: the living condition of people and the surrounding environment (natural, sociocultural, economic, institutional); the power structures that can influence the adoption of the systems produced by the project; the number of organisations and partners involved in the project and their role in pursuing results.

We have presented some essential notions of social impact assessment that can be used to evaluate AI projects, defined by mixed research-business groups, which need an impact analysis in sectors (let's make a hypothetical list), such as cultural heritage, smart cities, smart buildings, elderly care systems and educational systems.

5. Conclusions

In this chapter we distinguished two types of AI contributions: the functional one, due to the number of innovative systems it produces at the service of society, and the theoretical and methodological one that has an impact on many disciplinary areas. We examined the contribution of AI to the humanities, social and natural sciences. In particular to disciplines such as Linguistics, Cultural Heritage, Medicine and Education. We reported some critical issues of AI systems, we examined the problem of technological unpredictability, the unexpected results of AI systems, also discussing problems of regulation, opacity and prediction of the future of such systems. At the end of the chapter we reported some regulatory proposals of the AI systems commercialization and, we provided some elements for the impact analysis of these systems.

References

Bordoni, L., Ardissono, L., J.A. Barcelo and, A. C., de Gemmis, M., Gena, C., L. Iaquinta, P. L., Mele, F., Musto, C., Narducci, F., Semeraro, G., &

Sorgente, A., 2013, “The contribution of ai to understand cultural heritage”, *Intelligenza Artificiale*, vol. 7, no. 2, pp. 101-112.

Bordoni, L., Mele, F., & Sorgente, A., 2016, Artificial intelligence for cultural heritage, Cambridge Scholars Publishing, Cambridge.

Chomsky, N., 1975, *The logical structure of linguistic theory*, Springer, New York.

Cohen, J., 1988, “A view of the origins and development of prolog”, *Commun. ACM*, 31(1), 26–36. https://doi.org/10.1145/35043.35045.

Commission, E., 2021, Proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence - artificial intelligence act.

Cucchiara, R., 2021, *L’intelligenza non è artificiale: La rivoluzione tecnologica che sta già cambiando il mondo*, Mondadori, Milano.

Etica, I., & Consulting, S., 2016, *Le linee guida per la misurazione dell’impatto sociale,* http://www.improntaetica.org/wp-content/uploads/2016/06/Linee-Guida-Impatto_def.pdf (last access: 04/04/2022)

Group, A. H.-L. E., 2019, *Ethics guidelines for trustworthy ai*, European Commission, https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf (last access: 09/04/2022).

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D., 2018, “A survey of methods for explaining black box models”, *ACM computing surveys* (CSUR), 51(5), 1-42.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z., 2019, “Xai-explainable artificial intelligence”, *Science Robotics*, 4(37), eaay7120.

Hausser, R., & Hausser, R., 2001, *Foundations of computational linguistics*, Springer, Berlin.

Humphreys, P., 2009, “The philosophical novelty of computer simulation methods”, *Synthese*, 169(3), 615-626.

Interoperability, S., 2021, What is semantic interoperability [Available online]. https://www.igi-global.com/dictionary/challenges-interoperability-ecosystem/26340 (last access: 02/04/2022)

Kissinger, H., 2018, “How the enlightenment ends”, *The Atlantic*, 1.

Kowalski, R. A., & Sergot, M. J., 1986, “A logic-based calculus of events”, *New Generation Comput.*,4(1), 67–95.

Kunze, J. and T. Baker, 2007, "The Dublin Core Metadata Element Set", *RFC* 5013, DOI 10.17487/RFC5013, August 2007, <https://www.rfc-editor.org/info/rfc5013>.

Miller, R., & Shanahan, M., 2002, Some alternative formulations of the event calculus. In: Kakas A.C. & Sadri F. (Eds.), *Computational logic: Logic programming and beyond,* Springer, Berlin, pp. 452-490.

Mitkov, R., 2004, *The oxford handbook of computational linguistics*. Oxford University Press, Oxford.

Mueller, E., 2015, *Commonsense reasoning (second edition),* Morgan Kaufmann, Elsevier.

Numerico, T., 2021, *Big data e algoritmi: Prospettive critiche*, Carocci, Rome.

SAE International, 2018, *SAE international releases updated visual chart for its "levels of driving automation" standard for self-driving vehicles*, https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles (last access: 07/04/2022

Peyron, L., 2021, "Esserci, o non esserci? il dilemma dell'intelligenza artificiale", *MORALIA BLOG*. https://ilregno.it/moralia/blog/esserci-o-non-esserci-il-dilemma-dellintelligenza-artificiale-luca-peyron (last access: 08/04/2022).

Resler, A., Yeshurun, R., Natalio, F., & Giryes, R., 2021, "A deep-learning model for predictive archaeology and archaeological community detection", *Humanities and Social Sciences Communications,* 8(1), 1-10.

Riva, G., & Riva, E., 2020, "De-enigma: Multimodal human–robot interaction for teaching and expanding social imagination in autistic children", *Cyberpsychology, Behavior, and Social Networking,* 23(11), 806-807.

Rundo, F., Trenta, F., di Stallo, L., & Battiato, S., 2019, "Machine learning for quantitative finance applications: A survey", *Applied Sciences,* 9(24). https://doi.org/10.3390/app9245574.

Sadin, E., 2019, *Critica della ragione artificiale: Una difesa dell'umanità*. Luiss University Press, Rome.

Samuel, A. L., 1959, "Some studies in machine learning using the game of checkers", *IBM J. Res. Dev., 3*(3), 210-229. https://doi.org/10.1147/rd.33.0210

Semeraro, G., Ferilli, S., Girardi, E., Musto, C., & Marina Paolini, S. P., Piero Poccianti, 2021, *L'intelligenza artificiale per lo sviluppo sostenibile*. Cnr Edizioni, Rome.

Skeem, J., & Eno Louden, J., 2007, Assessment of evidence on the quality of the correctional offender management profiling for alternative sanctions (compas). *Unpublished report prepared for the California Department of*

Corrections and Rehabilitation. Available at: https://webfiles.uci.edu/skeem/Downloads.html (last access: 01/04/2022).

Stock, O., Strapparava, C., & Nijholt, A., 2002, *Twente workshop on language technology 20.*, ITC-IRST, Trento.

Stock, O., 1997, *Spatial and temporal reasoning*. Springer Science & Business Media.

Strubell, E., Ganesh, A., McCallum, A., 2019, Energy and Policy Considerations for Deep Learning in NLP. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Association for Computational Linguistics, Florence, Italy, pp. 3645-3650. https://doi.org/10.18653/v1/P19-1355

Tamburrini, G., 2020, *Etica delle macchine: Dilemmi morali per robotica e intelligenza artificiale*, Carocci, Rome.

Tim Berners-Lee, J. H., & Lassila, O., 2001, The semantic web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American 284*(5), 34-43.

Web, W.-S., 2021, Semantic web [Available on line]. https://www.w3.org/standards/semanticweb/

Yang, G.-Z., Cambias, J., Cleary, K., Daimler, E., Drake, J., Dupont, P. E., Hata, N., Kazanzides, P., Martel, S., Patel, R. V., *et al*., 2017, Medical robotics—regulatory, ethical, and legal considerations for increasing levels of autonomy. *Science Robotics, 2*(4), 1–2.

Yu, E., 1997, Towards modeling and reasoning support for early-phase requirements engineering. *Proceedings of the IEEE International Conference on Requirements Engineering*, 226-235. https://doi.org/10.1109/ISRE.1997.56687335.

The development of technology during the Anthropocene has affected science and the ways of "doing science". Nowadays, new technologies help scientists of several disciplines by facilitating knowledge and how to manage it, but also allow for collaborative science, the so-called "Social Science", where everyone can be a scientist and be involved in providing data and knowledge by using a computer or a smartphone without being a specialist. But is it really that simple? Actually, the daily and integrated use of different digital technologies and sharing platforms, such as social media, requires important reflections. Such reflections can lead to a rethinking of epistemologies and scientific paradigms, both in human geography and social sciences. This volume titled "Information Technologies and Social Media: New Scientific Methods for the Anthropocene" includes 10 chapters exploring some changes related to the way to do science with a multidisciplinary approach. From classroom experiences to the use of Citizen Science, from Artificial Intelligence use to how Social Media can help researchers, the book reflects on the ICT influence during the last few decades, exploring different cases, complementary perspectives and point of views.

Gaetano Sabato*, PhD in Tourism Sciences, is currently Researcher of Geography at the Department of Psychological, Pedagogical, Exercise and Training Sciences of the University of Palermo (Italy), where he teaches "Geography for Primary Education" at the Sciences of Primary Education master degree. He has published several scientific articles and a monograph: "Crociere e crocieristi. Itinerari, immaginari e narrazioni", published by Giappichelli, Turin 2018. Moreover, he is guest editor, with Leonardo Mercatanti, of two Special Issues of "AIMS Geosciences". His research focuses are on cultural geography and digital representations of the space, didactics, tourism, and perception of risk.*

Joan Rosselló *is an associate lecturer at the Open University of Catalonia. He holds a Physical Geography PhD, has published more than 20 papers in national and international journals and his research focuses are natural hazards, flash floods and precipitation, studying historical and contemporary events. He sits on the editorial board of the Geographies of Anthropocene book series, Physio-Géo Journal and the Journal of Flood Risk Management.*

OPEN ACCESS
PEER-REVIEWED SERIES
Geographies of the Anthropocene

ISBN 979-12-80064-36-3

IL Sileno Edizioni