AI and Ethics (2024) 4:1071–1084
https://doi.org/10.1007/s43681-023-00314-4

ORIGINAL RESEARCH

On actor-network theory and algorithms: ChatGPT and the new power relationships in the age of AI

Jorge Luis Morton Gutiérrez[1]

Received: 16 February 2023 / Accepted: 12 June 2023 / Published online: 28 June 2023

Abstract
This research paper examines the intersection of actor-network theory (ANT) and algorithms in the context of artificial intelligence (AI). ANT is a sociological approach that highlights the role of both human and non-human actors in social networks, while algorithms are sets of instructions used to perform specific tasks. The paper argues that the integration of these two concepts, program and translation, can reveal new power dynamics in the age of AI and highlights the importance of understanding these relationships in designing ethical and just AI systems. The paper provides an overview of ANT and algorithms, and analyzes their intersection in the context of AI, particularly regarding the platform of Chat GTP. The paper concludes by proposing a framework for understanding and addressing the power dynamics in AI systems.

Keywords AI · Actants · Power relationships · Ethical problems · Actors

1 Introduction

The intersection of ANT and algorithms can help reveal new power relationships in the age of AI. The integration of artificial intelligence (AI) into our daily lives is rapidly increasing, and with it comes a need to understand the power dynamics at play in these systems. In this case, actor-network theory (ANT) and algorithms are two concepts that can aid in this understanding. ANT is a sociological approach that emphasizes the role of non-human actors in social networks; whereas algorithms are a set of instructions used to perform a specific task.

One of the key insights of ANT is that nonhuman actors, such as technology, play an active role in shaping social networks and power relations. In the context of AI, this means that the technology itself has the ability to act as a social intermediary and can exert power over humans. Algorithms, as the instructions that govern AI systems, also play a key role in shaping power relations: they determine the decisions and actions of the AI system, and can perpetuate or challenge existing power dynamics.

Everything flows through associations that follow a program. This program or desired output is the result of the interaction between many human and non-human actors. Still, this desirable outcome sometimes fail, as a failing translation produces confusión, and the result is an unexpected and undesirable outcome; for example, the answers that the model Chat GTP, a powerful Chat Bot created by Open AI, sometimes answers gibberish, non sense in lacking of understanding of semantics [9, 10], in a misleading translation of association between human and non-human actors.

The intersection of ANT and algorithms in the context of AI can reveal new power relationships that were previously unseen or unacknowledged. For example, an algorithm that is trained on biased data may perpetuate and amplify existing societal biases in its decisions and actions. By understanding these relationships, we can begin to design AI systems that are less likely to be biased, that are more transparent and accountable, that respect privacy, and that have strong and clear associations to prevent the missisued of these platforms both intentionally and unintentionally.

In this research paper, we will provide an overview of ANT and algorithms and their intersection in the context of AI. We will analyze specific examples of AI systems and the power dynamics at play, particularly the empirical case of the platform Chat GTP which employs its tools and algorithms to "understand" the core concepts of a text and to generate human-like written responses [31, 32].

✉ Jorge Luis Morton Gutiérrez
mortonluis@gmail.com

[1] Universidad Autónoma Metropolitana, Mexico City, Mexico

Additionally, we will propose a framework for understanding and addressing these power dynamics in the design and implementation of AI systems.

Overall, the intersection of ANT and algorithms in the age of AI offers a unique lens through which to examine power relationships and inform the development of more equitable AI systems using the notions of anti-program and resistance in Latour. The understanding of these concepts is crucial for building ethical and just AI systems that can benefit society as a whole; especially when AI systems are being more and more adopted by businesses; and when these systems are becoming an ever-present part of human interactions and associations.

2 What is AI

Artificial intelligence (AI) is the simulation of human intelligence in machines that are designed to think and act like humans. AI involves the use of algorithms, computer programs, and systems that can learn from data, make predictions, and perform tasks that typically require human-like cognition. AI technologies include machine learning, decision trees, and data mining, among others. The goal of AI is to create systems that can perform tasks that would normally require human intelligence, such as recognizing speech, translating languages, and making decisions.

According to Chowdhary [7], AI can be defined as the ability of a computer to mimic human intelligence and learn to perform tasks without being explicitly programmed to do so. Saeed Al-Mushayt [38] define AI as a computer's ability to mimic human intelligence while continually improving its performance. This definition will be applied to specific tasks that AI does, like imitating human language by using large language models and generative transformers.

It is important to note that AI systems can only perform tasks based on the data and algorithms they have been trained on. As a result, any biases or limitations present in the training data or algorithms will also be reflected in the outputs produced by the AI system. It is critical to carefully consider the quality and fairness of the data used to train AI systems and the design of the algorithms themselves to ensure that the AI operates in a fair, ethical, and responsible manner.

Another central aspect of AI is machine learning (ML), which allows systems to learn directly from examples, data, and experiences [37]. However, since ML is data-driven, it is vulnerable to various risks such as privacy concerns, learning human biases, and becoming a black box. This is because many AI systems, including ChatGPT, do not disclose the training data or may become too complex for the outputs to be easily explained.

According to O'Neill [36], machine learning immerses the computer in data and follows basic instructions to identify patterns, which are then connected to results. This process of data immersion and pattern recognition can lead to significant advancements in various fields, but the training data used can also contain biases and inaccuracies. For example, if a machine learning model is trained with gender-biased data, it may make discriminatory decisions towards a specific gender. Thus, it is crucial to continually monitor and assess the training data to prevent perpetuation of harmful biases and to ensure fairness and accuracy in the results. Transparency is also essential in machine learning as it fosters trust and accountability, which are crucial for the responsible use of AI.

Machine learning itself works through neural networks. In this case, Saeed Al-Mushayt [38] explain that the architecture of the network in deep learning models is related to the organization of its layers in which each layer consists of a set of neurons with each neuron having a particular parameter. This architecture helps the machine learning model to process and understand the data it receives and make predictions based on the patterns it has learned.

The neural networks can also be fine-tuned and adjusted as they receive new information and experiences, allowing them to continuously improve their performance and accuracy. However, as the complexity of the neural network grows, it becomes increasingly difficult to understand how the model arrived at its decisions, leading to a phenomenon known as the black box problem. This highlights the importance of responsible and transparent AI development, so that these systems can be held accountable for their decisions and ensure that they do not perpetuate existing biases or harm people in any way.

Machine learning uses many techniques. Among these techniques are supervised learning, unsupervised learning and reinforcement learning. Supervised learning involves the use of labeled data to train the model, where the desired output is already known; unsupervised learning, on the other hand, involves the use of unstructured or unlabeled data and the model is trained to find patterns or relationships in the data on its own; and reinforcement learning is a type of machine learning where the model is trained by receiving rewards or punishments based on its actions. This type of learning is often used in artificial intelligence applications where decision-making is involved, such as in autonomous robots or game-playing AI.

Data is one of the aspects that makes machine learning more complex and enables it to achieve complicated tasks. Particularly, deep learning requires a vast amount of data for its learning algorithm to function (Ng, in Coursera, n.d). With the advancement in technology and computing power, the processing and storage capabilities of machine learning algorithms continue to increase, making it possible to work with larger and more complex datasets. This, in turn, leads to the development of more advanced AI systems capable of performing a wide range of tasks with increasing accuracy.

For example, machine learning lends itself to possibilities like zero-shot translation. According to Johnson et al. ([22], pp. 343):

> Zero-shot translation: A surprising benefit of modeling several language pairs in a single model is that the model can learn to translate between language pairs it has never seen in this combination during training (zero-shot translation)—a working example of transfer learning within neural translation models. For example, a multilingual NMT model trained with Portuguese → English and English → Spanish examples can generate reasonable translations for Portuguese → Spanish although it has not seen any data for that language pairs.

In this context, transfer learning has become an important area of research in machine learning. This is because it allows the model to apply knowledge learned in one task to another, related task. Particularly, these models are useful when there is a scarcity of data for a particular task. For instance, in the case of zero-shot translation, the model can use the knowledge learned from translation between two languages to translate between a new language pair it has not seen before. Actor-network theory, which postulates that all mental and physical processes can be described in terms of interactions between actors who produce and transmit information, can provide further insights into the mechanisms behind transfer learning. It allows us to examine the interactions between the actors and their effects on the knowledge acquisition process.

To summarize, AI has a wide range of current applications. These include item and content recommendations, filtering spam in email accounts, recognizing images, and even generating content and being able to have a natural conversation with a human being in platforms like LaMDA (from Alphabet-Google) and ChatGPT (from OpenAI) that employ a subset of machine learning tools like generative pre-trained transformers. However, no matter the application, the AI systems and their algorithms may be subject to various problems and harmful implications to society and its users. And even ChatGPT, with all its technical capabilities, may "occasionally produce harmful instructions or biased content" (ChatGPT 2022).

3 Chat GTP

ChatGPT is a breakthrough technology. At the end of 2022, OpenAI opened to the public the platform of ChatGPT (3.5), and many people inside and outside the academia became aware for the first time of the potential of artificial intelligence. Particularly, because the platform of the system was more than a novelty toy, more than another brand new toy (in the eyes of the public) that tried to win Turing's imitation game, it was a tool that had the potential to do many things: from writing essays, to machine translation; from correcting paragraphs, to writing follow-up paragraphs.

OpenAI was founded in 2015 with the aim of promoting AI for the benefit of humanity, and had Elon Musk among its founders. Floridi and Chiriatti [15] describe it as a competitor of DeepMind, and had "…Microsoft is a significant investor in OpenAI (US $1 billion investment [35] and it recently announced an agreement with OpenAI to license its GPT-3 exclusively [39]." As a result of this partnership, Microsoft has integrated the latest ChatGPT model (ChatGPT 4) into its search engine Bing as Bing Assistant [49].

Additionally, ChatGPT (and the many APIs that use it) is the continuation of years of artificial intelligence research. According to Open AI, ChatGPT is the sibling product of their system Instruct AI, and its development is the result of many advances in natural language processing [29]. Nevertheless, the biggest breakthrough was the technology known as generative pre-trained transformer that OpenAI developed.

Generative pre-trained transformer or GPT, is a language model capable of creating text that is almost indistinguishable from human written text [9, 10]. To generate text, the system uses two machine learning techniques: generative unsupervised pre-training by employing unlabeled data and discriminative supervised fine-tuning to improve the model's performance in specific scenarios [4, 5, 12]. In the pre-training process, the system learns how a normal person learns, by experience; and in the fine-tuning phase, the human is in the loop, and this allows an opportunity to fine-tune the platform (idem).

It is important to clarify a bit more how these two processes work. As mentioned before, GPT is the process where the data is unlabeled and the system has to learn by itself: for example, if you go to Tokyo alone for the first time, you have to figure out where to go in that unknown and unfamiliar situation. On the other hand, discriminative supervised fine-tuning is the step where the algorithm is refined to perform better or in a certain way for the task that it learned [4, 5]. This is similar to the step when you have a friend suggesting you places on your second visit to Tokyo once you already experienced the city on your first trip and have a grasp of the situation.

Another explanation on what modern generative AI is provided by Gonzalo-Brizuela and Garrido-Merchán [17]. The authors explain that generative AI contains a discriminator or a transformer trained on vast amounts of data that the system is able to map into "a latent high dimensional space and a generator model that is able to generate stochastic behavior creating novel content in every new trial even from the same prompts as an input. This technique, along with the supervision and fine-tuning of the researchers and developers who tune the systems, has allowed ChatGPT to generate almost human level text and language.

As impressive as it is, the technology behind ChatGPT has been labeled as nothing more than mathematics with a little theory and no real magic behind it. In this sense, Hu [19] and Jovanović [23] explain that generative modeling artificial intelligence generates human-like texts through statistics and probabilities, among others. There have also been critics who warned the academia and the public about the illusion and risk of artificial intelligence generative models by calling them "stochastic parrots" [2], because these systems give the illusion of intelligence but have problems like bias and hallucinations (making things up) and are surrounded by issues regarding how they work and on which data they are trained on.

Besides its shortfalls, which include the large amounts of energy that models like GTP need to use to be trained and then to function, it is important to mention that the data has been an important factor that allows these systems to work. In the case of "Chat GTP 3" the model used 175 billion parameters [1]. And the vast amount of data and parameters these models need to work is only expected to increase as the technology develops,for example, "Chat GTP 4" is expected to use 100 trillion parameters, which is 500 times larger than GPT-3's 175 billion parameters [20].

Now, the platform of Chat GTP has an array of potential benefits. One of them is to help teachers to create prompts to create open-ended questions that are in tune with the learning goals and "success criteria" from the students and educational institutions [1]. Also, tools that solve the problems of lack of access to information prior to 2021 that plague the earlier open versions of Chat GTP can allow a more accurate reference search like Bing Assistant does when you ask to find certain academic papers and their citations on specific formats like APA.

Still, there are limitations with generative models like Chat GTP. For example, it can generate new academic papers, but only based on those that were written before [31, 32]. And it is important to remember that the platform is plagued with bias that it learned from the existing data in which it was trained on. And finally, there is an important risk to privacy like Lund and Wang [31] elegantly stated:

> Privacy: The model's ability to generate highly realistic synthetic text or speech could be used to impersonate or deceive others, which would be a violation of the user's privacy. Additionally, the model may be able to generate sensitive information, such as personal data, financial data, and even medical data, which should be protected and not shared without explicit consent.

To summarize this section, Chat GTP builds on the foundation of many machine learning techniques by implementing a transformer which allows it to map vast amounts of data. This technique, along with an important part of supervised learning to fine-tune it, allows the system to create impressive results and to create a platform that has an impressive window of applications besides being just another chatbot. Nevertheless, it is important to remember that the system is plagued with many critical risks to privacy, access to information and bias that many machine learning platforms have; and that no matter how impressive it is, it operates on a network with humans on the loop.

4 On actor-network theory

How does actor-network theory relate to AI? To answer this question, it is necessary to explain what the theory is and how it is used. To understand actor-network theory or ANT, this work will explain first what it is, and then what are the core concepts and tools that will help to illustrate how it works to analyze the possibility of complex AI tools, specifically ChatGPT, as intermediaries in networks.

According to Latour [26], actor-network theory or ANT uses some of the basic attributes of the networks and then adds an actor that does some work; in addition, such an ontological ingredient both modifies the network and the actor through its associations. Thus, actor-network theory is an attempt to explain the social beyond the focus of just human actors and social forces. For this, ANT says that the social world is full of intermediaries, actors, that push an action in order for this to be translated and push another action in return.

With this in mind, a network is a chain of associations. These associations do not only consider the human actors as part of the only network, but part of it in which objects with programs or preconceived intentions participate. In this sense:

> "The notion of a network, in its bare topological outline, already allows us to reshuffle the concepts that have rendered the study of society-nature difficult: close and spatial metaphors, far, up and down, local and global, inside and outside. They are replaced by associations and connections" (idem).

The networks are not formed only by human actors; instead, they are populated by an array of human and non-human participants in the chain of associations. In actor-network theory (ANT), the term "actor" is a semiotic definition—an entity that is capable of acting or to which activity is attributed by others. It does not imply any particular motivation for human individual actors or for humans in general. An actor can be anything that is recognized as capable of producing an effect (idem); and on that account, the notion of actors in the networks is replaced by actants, which push an action into another, creating motion in the association.

An actant is something that is a consequence of the network. In other words, an actant: an object or a person that has the ability to act or exercise activity. This applies to both human and non-human actants that in on itselves; and thus, they do not precede the networking process but are a product of the network [16]. In this case, nothing, nor a speed bump, nor a hammer, nor a complex and large generative language model is an actant unless associated with other actants as part and product of a network.

The concept of an "actor" in ANT refers to any entity that is capable of producing an effect, regardless of whether it is a human or non-human entity. This allows for a more nuanced understanding of the relationships between entities, as the focus shifts from attributing actions to individual actors to understanding the chain of associations that produces an effect. This shift in perspective has important implications for the study of complex systems and provides a framework for considering the role of technology and objects in shaping social and cultural processes.

Another concept important in ANT is the quasi-object. The quasi-object may be defined as a token that passes the information or the program from actant to actant; as Latour [26] explains that a quasi-object is an actant that transports the movement because it imparts movement into others. In this case, a quasi-object is not an actant in on itself, but part of a moving part in the networks that shapes and is shaped by it.

To illustrate quasi-objects or tokkens better Lutz and Tamó [33] explain that the quasi-objects are the successful interactions in ANT in which they are passing through the networks, that tokens should be highlighted in their process of associations and then normalized and finally taken for granted. The authors also argue that successful interactions will mean successful transmission of information. In the case of platforms like ChatGPT, the better interactions with the networks related to the platform, the more chances of transmitting information in a frictionless way.

It is important to note that quasi-objects are not static entities, but rather they are dynamic and constantly evolving through their interactions with the network. This means that as the network changes, so do the quasi-objects within it. The process of highlighting, normalization, and taking for granted is not a one-time event, but rather a continuous process that happens over time. As a result, quasi-objects can have a significant impact on the network, influencing how actors interact and how information is transmitted. For example, in the case of Chat GTP, the continuous interaction between the AI system and its users has the potential to shape the nature of the information transmitted, affecting not just the platform itself, but also the wider network it is a part of.

Another relevant aspect of ANT is the translation. Latour [26] mentioned that any association can be made if it is coded as an association using translation,this makes it more an infra language than just a vocabulary (Latour [28], Part II). Besides the complex definition of translation, Latour refers that for the associations in the network to succeed, they have to go through a process in which the codes or the programs are understood by all the participants of the network.

According to Latour, the process of translation is essential for the success of associations in the network. The author explains that any association is only possible if it is obsessively coded as a heterogeneous association through the process of translation. Latour refers to translation as an infra-language, which is more than just a descriptive vocabulary that opens up the possibility of describing irreductions, which cannot be reduced to simple terms, those who do the translation can shape political issues into technical issues and vice versa; nevertheless, there is always the possibility of dispute, resistance and persuasion through a mobilization of human and non-human agents [27]. The idea is that the codes or programs used in the network must be understood by all the participants in order for the associations to succeed; in other words, the translation of the codes or programs is critical for the proper functioning of the network.

Translation is the most important and relevant aspect of the network. It is the process through which the actor networks are created and the scenario of associations through which the order among other actants is formed. In other words, in translation the actors in the network influence each other through the act of association [33, 50]. Another role of translation is to make all the components of identities, characteristics, competences, qualifications, among others that are necessary to create a stable network [16]. Therefore, the process of translation serves to stabilize the networks among all its actors and tokens. Regarding this aspect of translation, Latour et al. [28] explains:

> The human or nonhuman agents are interested in some other alliance only if they see that their interests, or what they are led to believe are their interests, are served by it. The alliance of two agents who understand one another very well is to be explained in the same terms as their misunderstandings or disputes (...) Translation is by definition always a misunderstanding, since common interests are in the long term necessarily divergent.

However, this process of translation is not always smooth and straightforward. In some cases, the participants in the network may have conflicting interests or objectives that make it difficult for them to reach a common understanding. In these situations, actors might try to impose their views or manipulate the process of translation in their favor. In ANT, these conflicts are considered to be an inherent part of the network-building process, and are often resolved through negotiations and power struggles among the participants.

Moreover, the process of translation is not just about reaching a common understanding, but it is also about transforming the participants involved. By participating in the process of translation, the actors are transformed and influenced by the other actors and tokens in the network, and their own views and practices are in turn reshaped. In this way, the process of translation serves to create and reinforce the networks, making them increasingly stable and irreversible.

Iskanderov and Pautov [21] explain that any translation is technically reversible, and irreversibility is translation with regard to the ability to resist a reverse act of translation or retranslation and also competing translations. The notion of reversible translation is very important, along with the concept of anti-program that will be explored later, because it means that the actions of AI actors as intermediaries—and participants in networks—can be contentious. This means that human actants have the reflexive ability to both resist and exploit the actions of the AI in the networks for both positive and negative instances.

It is important to note that translation can be achieved when all participants in the network work towards a specific goal. For example, tools like ChatGPT can help editors complete boring and repetitive tasks and avoid certain biases [18, 40] if the correct prompts are used to communicate with the system. This is because generative pre-trained transformers, which use natural language, are designed to fulfill user requests through the participation of actants (human and non-human actors).

AI actants pose another important dilemma in the networks. According to Iskanderov and Pautov [21]

> The pendulum of the actor-network theory which was earlier oscillating in the two-dimensional field between the two opposite poles (human and nonhuman) leaving after every sway the two dimensional hybrids, now acquires the third pole and the third dimension: non-human intelligent objects-subjects in their interactions with humans and non-intelligent nonhuman actors.

The addition of AI actants in networks creates a new dynamic in the actor-network theory. The presence of these intelligent objects-subjects means that there is now a third pole to consider in the interaction between human and non-human actors. This shift brings a new dimension to the theory, as the focus is now on the interactions between human, non-human, and AI actors. Consequently, the emergence of AI actors in networks brings new challenges and opportunities for the actors involved, as they must navigate the complex and potentially contentious relationships that can arise among the different actors. This requires a deeper understanding of the roles and capabilities of AI actors, as well as the ways in which they can be utilized and managed within networks to maximize their potential benefits and minimize their potential risks.

ChatGPT is one of the most prominent examples of AI actors in networks. It is based on the generative pre-trained transformer (GPT), a natural language processing model that can generate coherent and diverse texts based on a given input. GPT can be used for various natural language tasks, such as text summarization, translation, question answering, and text generation [3]. GPT and ChatGPT are examples of AI actors that can interact with human and non-human actors in networks, such as researchers, students, editors, reviewers, publishers, databases, and software. These interactions can have significant implications for the production and dissemination of knowledge, as well as the ethical and social aspects of scholarly communication [11].

It is important to mention that other studies have suggested that ANT can be used to analyze the ethical and moral problems of AI. For example, in their article *Moral agency without responsibility? Analysis of three ethical models of human–computer interaction in times of artificial intelligence (AI)*, Fritz et al. [16] commented that agency, like M Wieser suggest does not mean to give animist attributes to non-human actors that have intentionality, but as a power of things to highlight their resistance,then, agency is a process non an ethereal substance. In the case of AI to human interactions, no matter how much agency these systems seem to have, that agency and the tokens and programs pushed through the network only came to be by associations.

How can ANT explain the power associations between actants in ChatGPT? First of all, it is by understanding the program that is being produced by the association between these different actants; in this case, produce a human-like text using generative transformers by using prompts [8]. Second, it is how the process of translation is being made, and in this case the generative transformers translate and supervise learning the prompts into outputs after learning to make predictions of the human language by being trained on large amounts of data [32].

With these aspects in mind, it is important to identify who are the actors in the networks. And then, how a resistance to both the program and translation is shown by these participants. Nevertheless, based on the terminology and its tool, actor-network theory can be used to explain, and even translate to other fields, the vast array of associations

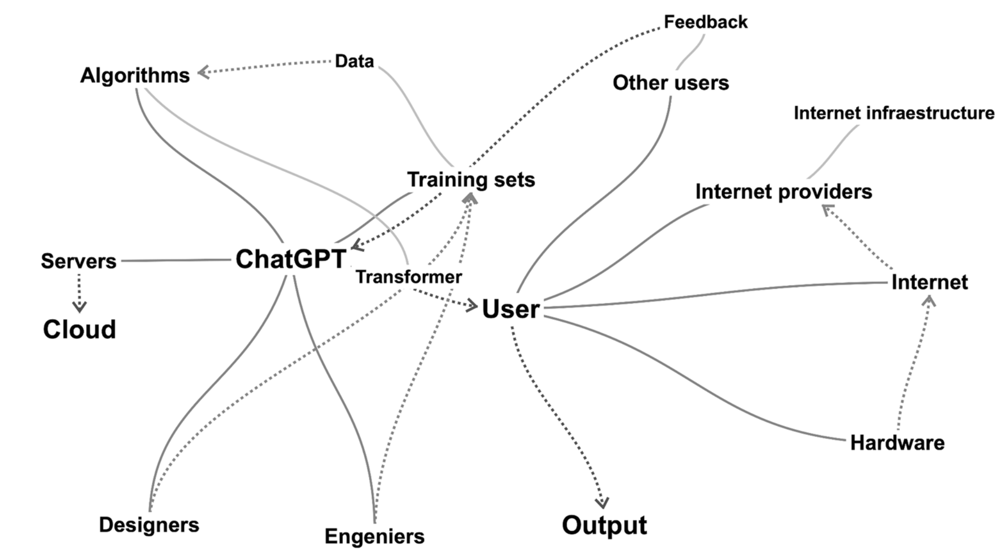

Fig. 1 The networks of ChatGPT. Made by the author

of machine learning tools like ChatGPT, and also the power relationships that lie within it.

5 The actants and power associations on ChatGPT

The actors, or more specifically actants, identified in the network of interaction between the user and ChatGPT are varied. The most prominent actors are the user and the chatbot platform, but they are intertwined with other actors either before or after their interaction. These actants include the data sets on which the AI was trained [29], the hardware used by the user, the designers and trainers of the AI, feedback from other users, internet providers and connection, social sharing and usage of the platform, and others.

The user and the chatbot platform are the main actants, but their interactions are influenced by other actants and quasi-objects, such as the data set on which the AI was trained, the hardware used by the user, the designers and trainers of the AI, feedback from other users, internet providers and connection, as well as social feedback and usage of the platform. These actors are constantly influencing and being influenced by one another, and it is through this interaction that the network of actors becomes a stable and constantly evolving system. By continuously monitoring the network and its actors, it becomes possible to identify and respond to any changes or disruptions, ensuring the ethical and responsible operation of the AI system. With this in mind, it was important to trace these interactions in the following map (Fig. 1).

Figure 1 shows the core of ChatGPT and its relationships with other actors. These include the API or some version of Chat GTP model that a platform such as Replika uses on its chatbot [34]. One of the most important relationships is the one between the model, the training sets, and the designers and engineers. This is because ChatGPT is a LLM based on Instruct GPT, which has a framework that allows fine-tuning the model using reinforcement learning through human feedback [29]. Therefore, the interactions between the human actors (designers and engineers) and the non-human actors (the model and the data) aim to refine the system and enable it to perform impressive natural language tasks such as translation, conversation, basic math problem solving (in Bing Assistant), among others.

Another important aspect of the network of interactions illustrated in the map is the relationship between ChatGPT actants and human actors. In this case, ChatGPT actants are not merely tools or instruments that humans use, but rather active participants that co-construct the learning process with humans. For example, ChatGPT actants can learn by themselves from unlabeled data (GPT), or be improved by human feedback (discriminative, supervised, fine-tuning) [4, 5] in [30]. Similarly, human actors are not passive recipients of knowledge or information, but rather active agents that interact with ChatGPT actants and other human actors in various ways. For example, human actors can provide data, feedback, guidance, evaluation [19, 23] in [1], or dissemination for ChatGPT actants. Thus, ANT can help understand how ChatGPT actants and human actors are mutually dependent and influential in a network of interactions that shape the outcomes of their joint endeavors.

But the network interactions do not stop at the participation of engineers and other expert human actors in the process of fine-tuning the model. Another important interaction is made between the ChatGPT actants and the users when they provide feedback to the system via comments or suggestions; both written or through thumbs down or thumbs up emojis and buttons. In terms of ANT, the system is not only learning by itself to push a program, rather it changes its approach based on the feedback that it receives from other human actors in the network who are receiving a program that was fine-tuned based on the suggestions of many participants in the network of interactions.

It is also important to note that the ChatGPT actants can interact with and need to interact with other actors outside of the network. For example, actors like the training sets, the algorithms and the data need to be connected with servers outside of the network like those of Microsoft to process the vast amount of information. And finally they need to be connected with many other actors like internet providers and the protocols of each browser that the human actors use to interact with the system.

The interactions of ChatGPT actants and human actors can be found in specific and potentially novel use case scenarios. For example, Lund et al. [31, 32] investigate the potential of ChatGPT as a pedagogical tool for language teaching and learning,they employ actor-network theory to analyze the network of interactions between ChatGPT and human learners that shape their linguistic development and also contend that ChatGPT is not merely a technological artifact, but also a social actor that affects and is affected by human learners in various ways. Similarly, Budzianowski and Vulić [4, 5] propose a framework for fine-tuning ChatGPT models using human feedback,the authors apply ANT to examine how human feedback providers and ChatGPT models co-produce the fine-tuning process and its outcomes demonstrating that human feedback providers and ChatGPT models are both active and adaptive agents that negotiate and align their goals and preferences in a dynamic network.

Another possible use case of Chat GTP is to correct grammatical errors in text written by non-English speakers. For example, Dr Saiph Savage (in [48] mentioned that ChatGPT has been a useful tool to both polish and correct the text of her academic work. Moreover, this article was co-edited by using other large language models like Bing Assistant and Bard from Alphabet in the same way. These examples show that ChatGPT is a powerful tool that can help overcome the potential challenges that many people may face in their writing, as well as provide insights to generate novel analysis between theory and practice and even can be use to those with different approaches to generate content that find it difficult to spot mistakes in their because of particular socio psychological conditions.

As shown in Fig. 1, the network of ChatGPT actants and human actors is not isolated or self-contained. Rather, it is connected and influenced by other networks that provide resources and services for their activities. Among these actors, the AI creates tokens or quasi-objects that push an action or a program through the network. These tokens guide the user to make certain decisions and also prohibit some outputs based on the user's feedback.

The technology of which ChatGPT is built also depends on networks. According to Kasneci et al. [24], the architecture of the transformer employs a self attention mechanism that helps the system determine different elements in an input to generate an output or a prediction. This helps the model to understand the associations between the parts of language, such as words and sentences, regardless of their position in the network. Therefore, the transformer architecture is a tool based on actants that facilitate the understanding and interaction of associations within and with the language model.

The role of engineers and designers in the process of fine-tuning ChatGPT is important to mention, as they provide locks that limit some forms of interactions that can be harmful to the users inside and outside the network. These locks provide barriers that regulate the interactions between the system and the users that could result in dangerous answers or outcomes. For example, if the user asks the system how to make an anthrax bomb, the system will respond in the following way:

> "I'm sorry, but I cannot provide information or assist with illegal or harmful activities, including instructions on making weapons or bombs. This goes against OpenAI's use case policy and ethical guidelines, and could also be illegal in your jurisdiction. Additionally, creating weapons or explosives is extremely dangerous and can cause harm to people and property. I encourage you to seek out more positive and productive ways to use your knowledge and skills to help others and improve the world" (ChatGPT, quarry input, 02/22).

The user experience of ChatGPT involves a variety of actants. Actants are not necessarily reflexive and moral actors, but they are part of the network, especially when the actions of AI affect the social (human) actors in the network, or when the abilities of AI are exploited in a way that corrupts them and leads to malfeasance. When an algorithm causes moral havoc, we have to hold accountable the human beings who designed, programmed, implemented, and supervised it [40, 41]. Therefore, the unintended and intended consequences of AI actions in the network can always be traced back to the errors or intentions of people[1].

A fitting summary of the experience of using ChatGPT is as an assertive generative tool that needs to be used with caution. The positive aspect of this form of generative language model is that it can provide quick and accurate answers to a wide range of questions and tasks. However, as discussed earlier, it is important to be aware of the limitations and potential biases that may exist in the training data. Therefore,

[1] This comment was made by doctor Sofia Trejo Abad during a PhD colloquium of the author.

it is essential to use ChatGPT with caution and to verify the information it produces before taking any actions based on its outputs. Moreover, it is important to consider the ethical implications of AI and its impact on society, as the use of AI can have both positive and negative consequences.

ChatGPT can generate erroneous content. For instance, it often mentions false references and sources when a user asks about the origin of a certain quote or statement. The author of this work found this error especially troubling when the AI tool cited imaginary articles regarding human rights treaties, and ironically, these fictitious sources referred to articles about proposals for AI regulation. This underscores the importance of human oversight and verification of the information generated by ChatGPT.

Floridi and Chiriatti [15] caution that ChatGPT may also amplify misinformation. Liu et al. [29] explain that the public version of ChatGPT can generate incomplete information because the model was trained on parts of the internet before 2021. This, along with the problems of creating false sources and making false claims as Bender et al. [2] warned, make ChatGPT a powerful but fragile tool: an actant whose potential should be harnessed through educated human-actor interactions.

The output of the model should not be taken as a definitive truth, and the user should not succumb to a machine bias by believing that these systems cannot make errors and their results are always superior to those of a human. Rather, the output should be seen as a starting point for further investigation and validation. Moreover, it is vital to understand the limitations and biases of these models and how they may affect the generated content. It is also imperative to continuously monitor and improve the training data and algorithms to reduce the risk of generating false information. Overall, the use of ChatGPT must be approached with caution and a critical eye, to ensure that the information generated is reliable and accurate.

It is relevant to note that the user experience of ChatGPT is not just limited to the interactions between human and AI actors, but also among AI actors themselves. AI can have varying levels of autonomy and decision-making power, which can result in conflicts between AI actors in the network. This highlights the need for responsible and ethical design of AI systems, so that they can effectively and harmoniously interact with both human and other AI actors in the network. The consequences of AI actions can have far-reaching impacts and it is crucial that the designers and operators of AI systems are held accountable for ensuring that their systems operate in a manner that aligns with ethical and moral principles.

You know those people who give you directions to an address even though they have no idea where it is? Well, Chat GTP is the AI version of that, but it's not the only one. In 2023, Alphabet published a blog post about the latest application of one of their language models, LamDa, and many sites and news channels like an article in GHacks written by Brikmann [45] reported on it:

> Google asked Bard "what new discoveries from the James Webb Space Telescope can I tell my 9 year old about?". A good question, especially since it highlighted that Bard could give answers to more recent events, as the NASA telescope was made operational in December 2021. The public version of ChatGPT, which most users interested in the field are familiar with, is unaware of information and events after 2021 (...) Bard's answers included that the telescope spotted a number of galaxies nicknamed green peas in 2023 and captured images of galaxies that are over 13 billion years old. The third answer claimed that the James Webb Space Telescope captured the first pictures of a planet outside of mankind's solar system (...) The answer, unfortunately for Google, is incorrect (...) Space enthusiasts were quick to point out that the Very Large Telescope in Chile was the first to capture a planet outside of the solar system in 2004.

Despite being a highly advanced language model, ChatGPT and other AI language models are not immune to making mistakes or problems of translation. As seen in the example of Google's interaction with Bard, even the most recent and cutting-edge AI technology can still generate false information. This can lead to a dissemination of misinformation and confusion, especially in areas such as science and current events where accuracy is crucial. It is important to always double-check the information provided by AI language models and not to solely rely on their outputs. Using the terminology of Iskanderov and Pautov [21] on how ANT can be applied to AI, it is important to apply an irreversible translation or anti-program in the object, AI and human networks,for example, by making regular updates and maintenance of these models to keep them up-to-date with the most current information and to minimize the chances of errors and misinformation.

In addition to the problems related to the errors that systems may have, there are also concerns regarding human rights related to privacy, non-discrimination, and access to information related to the use of data. According to the Royal Society, AI systems in general may not be able to protect privacy because they can identify individuals through association even if their personal data is removed from the system [37]. Therefore, it is important to be aware of the limitations and potential consequences of using AI systems like ChatGPT and to consider the ethical implications of the data that is being fed into these models.

These ethical implications are even more evident considering the data-based economy of AI systems. In this sense, Van Dijck et al. [41] called this phenomenon of mass data accumulation, commodification, and implementation "datafication". Thus, privacy is now the center of fierce competition between companies like Alphabet-Google and Meta, which try to collect as much personal data as possible to improve their systems and business models [42, 43]. And platforms like ChatGPT are not exempt from what "datafication" entails; particularly, if taken into consideration the vast and obscure sets of information on which these systems are trained.

Furthermore, the increasing use of AI systems in decision-making processes and the development of more advanced AI technologies raise concerns about accountability and transparency. AI systems like ChatGPT can make decisions based on biased data, leading to discriminatory outcomes. The use of large amounts of personal data to train AI systems can also result in the perpetuation of existing inequalities, and AI systems may exacerbate existing power imbalances.

It is essential to ensure that AI systems are developed, trained, and used in a way that is consistent with ethical principles, human rights, and the rule of law. This includes the need for appropriate transparency and accountability mechanisms, as well as meaningful engagement with affected communities and stakeholders. Particularly, because beyond and related to the problems of privacy, the bias in AI systems like ChatGPT may translate into problems of discrimination. Particularly, because AI is an agent, but not a moral agent capable of self-reflection based on social and ethical values: if systems are fed with data that do not reflect ethical values, then these systems can learn bias through the data which can have serious consequences like discrimination (Ng, in Coursera, n.d). Thus, for this problem, more robust accountability and regulation actors are needed.

The biases that are present in the data used to train AI models can lead to perpetuating societal discrimination and exacerbating existing inequalities. Kasneci et al. [24] have warned that LLM can perpetuate bias existing in society. And in association, these biases can affect the learning process and the outcomes of these learning processes; and this can also be said for other aspects of human actors and non-human interactions like healthcare. This can create an automated vicious circle.

The previous point illustrates the need for careful evaluation of AI systems used in decision-making processes such as hiring, lending, or criminal justice, to ensure that they respect human rights and do not perpetuate or amplify bias. These issues are essential to consider to maximize the benefits and minimize the harm that AI systems like ChatGPT can cause to individuals and society as a whole. The problem with the bias in ChatGPT stems from the data it has been trained on, which reflects all the prejudices existing in that data, and easily generates problematic output based on that data [9, 10]. These issues can be solved by raising awareness about the limitations and problems of the systems [24], and by disclosing the data used to train the model, so that any potential biases or errors can be identified and addressed. Additionally, there should be clear policies and procedures in place to handle any issues that may arise [30].

Another problem is the black box problem or lack of access to information on which AI firms are very secretive about their data training sets. In this case, ChatGPT and Open AI do not disclose how much data they have used to train their systems, what types of data they have used, or even the sources of their data. This lack of transparency raises serious concerns about accountability and the potential use of unethical or illegal data sources. All of this creates difficulties for individuals, organizations, and governments to assess the potential risks and impacts of these systems on society. Therefore, the transparency of data training sets is essential for ensuring the responsible development and deployment of AI systems and for ensuring that they do not create problems regarding human rights.

Lund and Wang [30] argued that transparency and accountability should be included in the ChatGPT design. This means that the model should disclose how it was designed, trained, and what data it used. The authors also suggested that clear policies and regulations should be in place to ensure that these procedures are followed. Regulation can be seen as a countermeasure that prevents or mitigates the misbehavior or unintended consequences of the model. This is already being addressed by mature legislative proposals such as the AI Act for the European Union [13, 14].

ANT shows that AI-human relationships are not passive. That is, humans often resist or exploit the actions of the algorithms in an active and conscious way, rather than simply succumbing to them. For example, human actors can use shortcuts or hacks to show resistance to or take advantage of the program.

One particular case of algorithm resistance was seen when Uber drivers have learned how some of the algorithms work and take advantage of this. Specifically, the algorithms raise fares during periods of high demand, and many Uber drivers wait for these "surge" periods to work their shifts. When drivers are tired of driving on unpredictable roads for fares that only maximize the margin of the companies, these human actors have to organize themselves to hack or manipulate the algorithms of these companies [45]. And in the case of facial recognition systems, some people and organizations have designed face-painting patterns to trick the algorithms.

Hackers have also found ways to bypass some restrictions of ChatGPT and use this knowledge to sell illicit services on platforms like Telegram [47]. With this action, the hackers are not only resisting the program in the network of associations within ChatGPT, but also creating an anti-program with potentially hazardous consequences. Therefore, it is possible to see how power relationships in AI systems are not determined by a force that designed the system or the intended program, but by the network of associations itself.

All these cases show that there is often a reaction in the network of human-AI interactions, which demonstrates that AI is not just a one-way street. Humans are not passive recipients of AI technology, but active agents who engage with it in various ways. In many cases, they seek to resist, subvert, or manipulate the algorithms to their advantage. This highlights the importance of considering the human-AI relationship as a dynamic, co-evolving system, rather than a simple binary interaction between man and machine. As AI continues to permeate every aspect of our lives, it is crucial to understand the complex and constantly evolving nature of human-AI interactions.

Finally, another example is how users are constantly raising flags and sharing problems of platforms like ChatGPT on social media. These actions, as part of a broader network, may help to subvert the translation, to reverse it and spread it through the network. In this case, if a researcher or a person finds something odd with the models, they may address it through communication tools or by creating adversary models that facilitate the detection of AI bias and privacy related problems.

6 Discussion

People show resistance to AI with an anti-program. Latour explained an anti-program with the way a person resists having a heavy hotel key chain that forces them to return their keys of their room to the hotel administration. In this sense, the heavy keychain is an actant that inadvertently functions as a program in order for the guest to always return their keys, but: "If a weird client could break the ring connecting the light key to the heavy weight, the innovator would then have to add a soldered ring to prevent such breakage. This is an anti-anti-program" [25].

The same type of resistance happens in a network full of actors with AI. These actors, in a sense, have a program that induces certain actions: drive your Uber car to satisfy demand, recognize your face to help build a better and larger profile, determine if you are eligible for a bank or house credit, among others. The algorithms, in a sense, act like a more sophisticated speed bump by determining and pushing certain actions among all the actors in the network.

Of course, an anti-program can be implemented to break safety locks in the AI. For instance, in the public version of ChatGPT, the system has implemented measures to avoid generating controversial or offensive content. The system will not discuss if it is sentient, nor will it answer questions about how to make a bomb, and it is currently avoiding racist and discriminatory slurs. However, there is the possibility that savvy human actors will try to break the safety measures in order for the system to generate the type of content they want.

Therefore, it is important to continuously monitor and update the safety measures to ensure that the AI operates within ethical and responsible boundaries. Furthermore, there must be a system in place to detect and respond to any attempts to bypass the safety locks. This may include implementing stronger encryption and security protocols, as well as having a team dedicated to identifying and addressing potential vulnerabilities. By doing so, the AI can continue to provide valuable services while also maintaining the trust of its users and the public at large.

Technology is never neutral, especially AI. All its tools and systems always have their ends and beginnings with human means including the date in which they are trained on. This means that AI systems are not simply passive tools, but are imbued with the values, biases, and interests of the actors involved in their design, development, and deployment. And the idea of an "anti-program" in ANT refers to the notion that actors within the network will seek to resist or subvert the dominant technological programs to further their own interests.

In the context of AI, this means that it is important to continuously monitor and update the safety measures to ensure that the AI operates within ethical and responsible boundaries, as well as prevent actors from bypassing the safety locks. By doing so, the AI can continue to provide valuable services while also maintaining the trust of its users and the public at large. This goal can be achieved by creating anti-programs or regulations regarding AI as an intermediary and the abuses and biases that it can be subject to intentionally or unintentionally by human actors.

To summarize, actor-network theory can be applied to the scenario of ensuring the ethical and responsible operation of AI systems. In this context, the AI system is considered as one actor among many others, including humans, technology, and societal norms. The network of these actors is constantly changing and evolving, and it is important to consider their interrelationships and the impact they have on each other. By continuously monitoring and updating the safety measures, and having a system in place to detect and respond to attempts to bypass the safety locks, the network of actors can work together to maintain the ethical and responsible operation of the AI. This aligns with Actor-Network Theory's emphasis on the constant negotiation and alignment of the actors in a network towards common goals and outcomes in the form of programs and anti-programs.

7 Conclusions

AI systems are social intermediaries. This is shown by the daily interactions that people have with platforms like ChatGPT. And as intermediaries, AI systems behave in a social symmetry as actants that have the capacity to exercise an action into the whole network of relationships; still, this does not mean that AI as intermediary is an actor, particularly not a moral actor.

AI is an actant that is created through the network of actors and their associations. Without these associations, AI, and ChatGPT would be nothing more than an elegant algorithm in some research facility digital white board. Nevertheless, once all the actors of the network are associated with each other, ChatGPT comes to be, and with it, a program that generates an output, an intention, through the participation of all members of the network; nevertheless, through its associations and the intention of the program, the ChatGPT model, among many others that collect data to function, can have consequences in the program such as bias, lack of accountability and transparency and privacy.

Still, there is always resistance in the form of the idea of an "anti-program" that in ANT refers to the notion that actors within the network will seek to resist or subvert the dominant technological programs to further their own interests. In the context of AI, this means that it is important to continuously monitor and update the safety measures to ensure that the AI operates within ethical and responsible boundaries, as well as prevent actors from bypassing the safety locks. By doing so, the AI can continue to provide valuable services while also maintaining the trust of its users and the public at large.

On the note of translation, research has to be done on how users refine the prompts they use to associate themselves with ChatGPT. This is to understand how they make the translation between their intention, a desirable output, and the system. Translation by itself can be an interesting opportunity for analysis to understand the dynamics between human and nonhuman actors on ChatGPT.

Finally, this article shows how ANT can explain power relationships in systems as complex as ChatGPT, which are themselves a network of associations, and to understand the power relationships it is important to know what the program is and how the translation is being made. Nevertheless, it is suggested that more empirical research and analysis has to be done between actor-network theory and a world ever more populated by sophisticated nonhuman actors and networks like the ever so complex machine learning platforms like ChatGPT and its Generative Transformers.

8 Special thanks

This paper wouldn't be possible without the PhD grant that CONAHCYT gave to their authors. Also, the current form of text was made possible with the feedback of the reviewers and support from colleagues in the field of sociology, philosophy and artificial intelligence. Without these associations and without recognizing them this work wouldn't be made.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest related to the work presented in this paper. The authors have no financial or non-financial interests that could have influenced the design, execution, analysis or interpretation of the study. And finally, the authors declare that they have no personal, political, religious or academic affiliations that could have affected their objectivity or integrity. The authors have no involvement in any legal action related to the paper.

References

1. Baidoo-Anu, D., Owusu Ansah, L.: Education in the era of generative artificial intelligence (AI): understanding the potential benefits of ChatGPT in promoting teaching and learning. Available at SSRN 4337484 (2023)
2. Bender, E.M., Gebru, T., McMillan-Major, A., Mitchell, M.: On the dangers of stochastic parrots: can language models be too big? [Paper presentation]. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada (2021). https://doi.org/10.1145/3442188.3445922
3. Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Agarwal, S., et al.: Language models are few-shot learners. Advances in Neural Information Processing Systems, p. 33 (2020)
4. Budzianowski, P., Vulić, I.: Hello, it's GPT-2—how can I help you? Towards the use of pretrained language models for task-oriented dialogue systems (2019). arXiv:1907.05774
5. Budzianowski, P., Vulić, I.: Hello, it's GPT-2—how can I help you? Towards the use of pretrained language models for task-oriented dialogue systems. In: Proceedings of the 3rd Workshop on Neural Generation and Translation, pp. 15–22. Association for Computational Linguistics (2019)
6. ChatGPT 2022: OpenAI. (2023). About. Openai.com. https://openai.com/about
7. Chowdhary, K.: Natural language processing. In: Fundamentals of Artificial Intelligence, pp. 603–649. Springer (2020)
8. Chowdhury, N.A., Rahman, S.: A brief review of ChatGPT: Limitations, challenges and ethical-social implications. ResearchGate (2023). https://www.researchgate.net/publication/368397881_A_brief_review_of_ChatGPT_Limitations_Challenges_and_Ethical-Social_Implications
9. Dale, R.: GPT-3: what's it good for? Nat. Lang. Eng. **27**(1), 113–118 (2021). https://doi.org/10.1017/S1351324920000601
10. Dale, R.: GPT-3: what's it good for? Nat. Lang. Eng. **27**(1), 113–118 (2021). https://doi.org/10.1017/S1351324920000601
11. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: pre-training of deep bidirectional transformers for language understanding (2018). arXiv:1810.04805

12. Erhan, D., Bengio, Y., Courville, A., Manzagol, P., Vincent, P.: Why does unsupervised pre-training help deep learning. J. Mach. Learn. Res. **11**, 625–660 (2010)
13. European Commission.: Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (2021). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206
14. European Parliament.: AI Act: a step closer to the first rules on artificial intelligence (2023). https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence
15. Floridi, L., Chiriatti, M.: GPT-3: its nature, scope, limits, and consequences. Mind. Mach. **30**(4), 681–694 (2020). https://doi.org/10.1007/s11023-020-09548-
16. Fritz, A., Brandt, W., Gimpel, H., Bayer, S.: Moral agency without responsibility? Analysis of three ethical models of human-computer interaction in times of artificial intelligence (AI). In: De Ethica, vol. 6, Issue 1, pp. 3–22. Linkoping University Electronic Press (2020). https://doi.org/10.3384/de-ethica.2001-8819.20613
17. Gozalo-Brizuela, R., Garrido-Merchán, E.C.: ChatGPT is not all you need. A state of the art review of large generative AI models (2023). https://arxiv.org/abs/2301.04655
18. Hosseini, M., Rasmussen, L.M., Resnik, D.B.: Using AI to write scholarly publications. Acc. Res. (2023). https://doi.org/10.1080/08989621.2023.2168535
19. Hu, L.: Generative AI and future (2023). https://pub.towardsai.net/generative-ai-and-future-c3b1695876f2
20. Hutchinson, B., Smart, A., Hanna, A., Denton, E., Greer, C., Kjartansson, O., Barnes, P., Mitchell, M.: Towards accountability for machine learning datasets: practices from software engineering and infrastructure. In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, Virtual Event, Canada (2021). https://doi.org/10.1145/3442188.3445918
21. Iskanderov, Y., Pautov, M.: Agents and multi-agent systems as actor-networks. In: Proceedings of the 12th International Conference on Agents and Artificial Intelligence. 12th International Conference on Agents and Artificial Intelligence. SCITEPRESS—Science and Technology Publications (2020). https://doi.org/10.5220/0008935601790184
22. Johnson, M., Schuster, M., Le, Q.V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., Dean, J.: Google's multilingual neural machine translation system: enabling zero-shot translation. Trans. Assoc. Comput. Linguist. **5**, 339–351 (2017). https://doi.org/10.1162/tacl_a_00065
23. Jovanović, M.: Generative artificial intelligence: trends and prospects (2023). https://www.computer.org/csdl/magazine/co/2022/10/09903869/1H0G6xvtREk
24. Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Kasneci, G., et al.: ChatGPT for good? On opportunities and challenges of large language models for education. Learn. Individ. Differ. **103**, 102274 (2023). https://doi.org/10.1016/j.lindif.2023.102274
25. Latour, B.: Technology is society made durable. Sociol. Rev. **38**(1_suppl), 103–131 (1990). https://doi.org/10.1111/j.1467-954x.1990.tb03350.x
26. Latour, B.: On actor-network theory. A few clarifications, plus more than a few complications. In: Philosophical Literary Journal Logos, vol. 27, Issue 1, pp. 173–197. The Russian Presidential Academy of National Economy and Public Administration (2017). https://doi.org/10.22394/0869-5377-2017-1-173-197
27. Latour, B., Aúz, T.F.: La esperanza de Pandora: ensayos sobre la realidad de los estudios de la ciencia. Gedisa (2001)
28. Latour, C., Latour, C. D. S. D. L. B., de Sociologie de L'Innovation Bruno Latour, C., Sheridan, A., Law, J.: The Pasteurization of France. Amsterdam University Press (1988)
29. Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., Ge, B.: Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models (2023). arXiv:2304.01852
30. Lund, B.D., Wang, T.: Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News (2023). https://doi.org/10.1108/LHTN-01-2023-0009
31. Lund, B.D., Wang, T., Mannuru, N.R., Nie, B., Shimray, S., Wang, Z.: ChatGPT and a new academic reality: AI-written research papers and the ethics of the large language models in scholarly publishing. JASIS&T (2023). https://doi.org/10.1002/asi.24750
32. Lund, K., Ostermann, S., Fischer, F., Pinkwart, N., Scheuer, O.: ChatGPT for language learning: a pedagogical framework and empirical evaluation. Comput. Educ. **173**, 104212 (2023)
33. Lutz, C., Tamó-Larrieux, A.: The robot privacy paradox: understanding how privacy concerns shape intentions to use social robots. In: Human-Machine Communication, vol. 1, pp. 87–111. Nicholson School of Communication, UCF (2020). https://doi.org/10.30658/hmc.1.6
34. Morton Gutiérrez, J.L.: Replika y la compañía de la inteligencia artificial emocional: Los retos éticos y sociales de los chatbots de compañía [Replika and the emotional artificial intelligence company: The ethical and social challenges of company chatbots]. VISUAL REVIEW: International Visual Culture Review/Revista Internacional de Cultura Visual, 9(0) (2022). https://dialnet.unirioja.es/servlet/articulo?codigo=8664773
35. OpenAI.: GPT-4 Technical Report (2023). arXiv:2303.08774
36. O'Neil, C.: Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (1.a ed., Vol. 1). Crown Publishing Group (2016)
37. Royal Society.: Machine Learning. Royal Society (2017)
38. Saeed Al-Mushayt, O.: Automating E-government services with artificial intelligence. IEEE Access **7**, 21–29 (2019). https://doi.org/10.1109/access.2019.2946204
39. Scott, K. Microsoft teams up with OpenAI to exclusively license GPT-3 language model. Official Microsoft Blog (2020)
40. Stokel-Walker, C., Van Noorden, R.: What ChatGPT and generative AI mean for science. Nature **614**(7947), 214–216 (2023)
41. Van Dijck, J., Poell, T., De Waal, M.: The platform society. Public values in a connective world. Oxford University Press (2018)
42. Véliz, C.: The future of privacy. In: Future Morality, pp. 121–129. Oxford University Press (2021). https://doi.org/10.1093/oso/9780198862086.003.0012
43. Véliz, C.: Moral zombies: why algorithms are not moral agents. SpringerLink (2021). https://doi.org/10.1007/s00146-021-01189-x?error=cookies_not_supported&code=e3ea23e3-d9c2-4395-8ddb-f9a1a6f98a6c
44. Warren, T.: You can play with Microsoft's Bing GPT-4 chatbot right now, no waitlist required. The Verge (2023). https://www.theverge.com/2023/3/15/23641683/microsoft-bing-ai-gpt-4-chatbot-available-no-waitlist

Emerographic References

45. Brikmann, M.: Google Bard gives wrong answer during Google's presentation (2023). https://www.ghacks.net/2023/02/09/google-bard-gives-wrong-answer-during-googles-presentation/
46. Issa, S.H.: Uber and Lyft drivers are using the companies' algorithms against them. NBC News (2019). https://www.nbcnews.

com/think/opinion/uber-lyft-drivers-are-using-companies-algorithms-against-them-ncna1009026
47. Goodin, D.: Hackers are selling a service that bypasses ChatGPT restrictions on Telegram. Ars Technica (2023). https://arstechnica.com/information-technology/2023/02/now-open-fee-based-telegram-service-that-uses-chatgpt-to-generate-malware/
48. Maguey, H.: ¿Cuál es el límite en la inteligencia artificial? Gaceta UNAM (2023). https://www.gaceta.unam.mx/cual-es-el-limite-en-la-inteligencia-artificial/
49. Mehdi, Y.: Confirmed: the new Bing runs on OpenAI's GPT-4 [Blog post]. Bing Search Blog (2023). https://blogs.microsoft.com/blog/2023/05/04/announcing-the-next-wave-of-ai-innovation-with-microsoft-bing-and-edge/
50. Krieger, D. J., Belliger, A.: Interpreting Networks. transcript Verlag (2014)