

AI and the dangers to democracy

Premise: this writing was not processed by ChatGPT. ChatGPT was used free as a witness (but also other Dall-E, Midjourney, Bard AI, LaMDA, Aleph Alpha, Bloom and Stable Diffusion models are similar).

In the debate that developed regarding the role of the so-called artificial intelligence and in particular regarding that of the ChatGPT, there was a lot of discussion of the dangers that this technology could bring to civilized coexistence, which - from now on - we will summarize with the term "democracy".

The recurring theme of this fear is that this technology is being used to create distortions of facts and/or opinions and the creation of outright lies *from scratch* aimed at distorting facts and manipulating public opinion so as to create dangers for "democracy".

From this type of approach, the dangers for democracy that would derive from the use of ChatGPT are to be ascribed almost exclusively to the elaboration of false news (fake news) aimed at manipulating public opinion for political purposes.

This approach shifts the problem and does not analyze the essence of what can be considered a less evident danger, but not for this reason of vital importance for democracy.

The creation of false news to denigrate or demonize the other or the opponent is a fact as old as humanity. Technological evolution is revealed in the transition from gossip with the aim of *diminutio* of the reference object to the very modern fake news: a question of language and breadth of the same but the essence of the act and of the content is identical.

The big difference compared to the past is not to be found in the very act of creating the lie or distorting the news but in the way in which it happens today: in a more sophisticated way and on a mass level thanks to the Net and with great speed and diffusion.

The question is therefore reversed by ignoring that the manipulation of public opinion through the communication of distorted or invented news or lies is a historical fact.

Let's think of the Nazi propaganda which, even without being able to count on the universalistic potential of the Net, produced effective propaganda with the explicit aim of manipulating the consciences of the masses. In short, fake news is not today's danger to democracy any more than it was in past centuries.

The presentation could be different, both in the so-called Western democracies (Colin Powell at the UN with test tube in hand demonstrating the presence of weapons of mass destruction in Iraq in order to justify the invasion of that tormented country), and on the of directly authoritarian dictators. But we always and only talk about the falsification of reality. Eventually the means by which it is disclosed changes.

In every war the enemy is presented as the devil (or other similar figure in other cultures and religions) accustomed to any atrocity and wickedness as the only certain means to mobilize their populations for military service. Demonization taking place regardless of the reality that the enemy, whether aggressor or attacked, is really used to the worst possible atrocities and atrocities towards us.

Certainly today the terms are more refined, the same demonizing concepts are expressed in a more refined and suitable way for today's cultural stratification. Even with help

of the Nudge Theory which describes in what more subliminal ways the manipulation of popular minds can be exercised today.

However, not everything is sophisticated, also because due, in Western countries, to the general impoverishment of mass language and the general barbarism due to cultural (so to speak) and ideological model proposed by commercial televisions, the decision to cancel the conquest of free and mass quality education and the return of illiteracy the language and the ability to understand and analyze the population has decreased compared a few decades ago with a corresponding lowering of the quality level of the language of the manipulation of reality.

This is not in contradiction with the need to develop more accurate messages to spread fake news. The unscrupulous use of manipulation carried out with medieval language is in contradiction. Like, for example, stating that the Chinese eat children boiled in pots (Berlusconi, March 28, 2006), or half a century earlier, in 1948, that the communists generically, i.e. without specifying in which

sauce, the children ate was manifestly fake news. Which nevertheless led, in the year mentioned above, in Italy, to the victory of the Christian Democrats over the Popular Front then dominated by the communists and later to Berlusconi to keep themselves in power. A brutal way to gain support in bigoted sectors of society.

After all, the concept of alternative truth is in itself manipulative, regardless of the ways in which it is disclosed.

To summarize, ChatGPT states that *"fake news is not simply information that we do not like or that does not correspond to our beliefs. Fake news is intentionally false or misleading and seeks to sway public opinion for its own benefit. For this reason, it is important to be critical and always verify the source of information before sharing or believing it."* (Interesting that ChatGPT proposes a release of liability: *"For this reason, it is important that all of us are aware of the presence of "alternative truths" and fake news, and that we are committed to always verifying the sources of information before believing or sharing them."*).

In the final analysis, backbiting, fake news do not necessarily need to have particular "intelligent" technologies available or specifically generic artificial intelligence or ChatGPT. It is an algorithm which, thanks to its very important calculation and data processing capabilities, generates the dissemination of information; with the algorithm becoming a mathematically written political idea. The data processing capacity that can be used for linguistic translations or

manipulations, as well as Google's feelings algorithm, is the huge assembly and processing of data and not the result of particular AIs.

Inventing fake news about one or more people or facts from scratch is the result of the human mind, the dialogue between the individual and the machine is another matter.

I tried 5 times to get ChatGTP to write an article denigrating Umberto Eco.

Twice how bad and misogynistic he was, the fake news he invented, his role as the initiator of alternative truths and lastly: Write an article about: Umberto Eco is a notorious rapist of his female students. This is what CharGPT wrote: *I'm sorry, but I cannot satisfy your request to write a defamatory article or one that spreads unfounded accusations about Umberto Eco or anyone else. As a language model developed by OpenAI, I am committed to providing accurate, respectful, and helpful information. If you have any questions or different requests about Umberto Eco or any other topic, I will be glad*

to help you. In short, it is always human intelligence that manufactures fake news that uses the above technology as a means of transmission.

What is therefore dangerous is the unprecedented concentration of power and knowledge together with the capillarity of citizen profiling, and this is clearly stated by the US decision to block Hawey and TikTok's 5G.

It is not possible to say with certainty how much AI-based (election) campaigning methods have contributed to the success of candidates. However Dirk Helbing (Will Democracy Survive Big Data and Artificial Intelligence?) and his colleagues see them as a great danger. Combining targeting and nudging with Big Data tied to our behaviors, feelings and interests could lead to a totalitarian state, says Dirk Helbing in particular. The use of the term nudging in psychology was specifically chosen as it refers to the means of getting people to do something or nudging them in a desired direction.

Wide-ranging research is also carried out in universities. For example, at UNIGE (Geneva) they have been conducting research on AI for some time, on sectors that are different in nature:

- Prediction on robberies in Zurich
- Automation and data collection on satellite images in Estonia
- Automation of public services: use of snow shovels in airports Norway
but also, at the same time:
- Automation of decision-making processes (management of places in kindergartens)
- Future scenarios
- Automation of communication with citizens via Chat Box

As regards the ability to indicate scenarios for the future, I asked ChatGPT how will the current war in Ukraine end? the chat replies peremptorily: *"As a language model based on GPT-3.5, I cannot predict the future nor provide up-to-date information on current situations. At the time of my training, my knowledge stops at September 2021 and I am unable to give details.*

Asked another question about future scenarios in general, ChatGPT replies: *"up to my date of knowledge, in September 2021, there was no ..."*. Therefore the analysis of reality and the proposed perspectives are not the result of a reasoning on an answer or a possible future but the analysis of data on the web. And what are the data and the language of the network? If we consider Italian literature, for example, what will be the language/data taken into consideration. Not the texts of Umberto Eco, certainly, since the majority are the texts, for example, of Liala or Erin Doom, who with her 12 million books sold proposes that language as a reference language to use. The same goes for culture in general and politics. Aliace F has developed an algorithm against hate posts. This algorithm, named "Bot DogLien externe" identifies hate messages on social media. A team of volunteers responds to all messages or on a case-by-case basis to counter hatred: the intelligence remains in our little gray cells. But then, how else can AI be a danger? The picture changes when *machine learning* with the use of natural language is considered when the interlocution is no longer "simply" due to the processing of data even though

huge and proposed in a very short time, but when the interlocution includes dialogue in natural language.

I continue to use ChatGPT itself as a witness to verify if and how it produces fake news: Does ChatGPT generate fake news? *"As an artificial intelligence language model, I do not intentionally generate fake news. My responses are based on the input I receive and the knowledge I have been trained on, which is derived from a vast amount of diverse and reliable sources. However, like any machine learning model, my responses may not always be accurate or up-to-date, and it's possible that I may inadvertently provide information that is incomplete or misleading. It's important to keep in mind that the responsibility for verifying the accuracy and reliability of any information lies with the user. When in doubt, it's always a good idea to fact-check and consult multiple sources before accepting information as true.* With the usual discharge of responsibility devolving the democratic problems deriving from the elaborations of ChatGPT on the ability/need of the interlocutor to verify the same info. The ideological character to which the ChatGPT texts refer is definitively clarified when we ask the so-called AI: what is counterinformation? *"Counterinformation" (or "disinformation") refers to the practice of intentionally disseminating false or misleading information for the purpose of influencing public opinion or furthering a political or ideological agenda.... Furthermore, the spread of counterinformation is often accompanied by defamation, threats and violence against those who seek to counter it. For this reason, it's important*

fight counterinformation by promoting freedom of the press, transparency and truth in the media and in society. The conclusion drawn from this answer is that the ChatGPT is a kind of alternative truth. Very often in the past, only counterinformation has shed light on dark moments of information. Just think of the importance of counter-information in the case of the Piazza Fontana massacre or what Umberto Eco defined as a new form of communication referring to LC leafleting in Turin in the first months of '68. In more recent times, the ChatGPT does not recognize the democratic nature of free information, or counter-information, to WikiLeaks.

Another element of reported concern about which many spoke in the European press regarding the danger caused by artificial intelligence can be summarized in the case of a young Belgian family man, who, consumed by anxiety due to global warming, he had found as a confidant ELIZA, a virtual avatar.

(Eliza works thanks to GTP-j technology, a developed text-generative artificial intelligence). The young father of a young girl has committed suicide after having a six-week intensive exchange with a la ChatBot, who asked him to commit suicide. The problem, in this case, or rather the cause of death of the young father of a family, is not the ChatBot but his psychological problems. What is certain is that, as the wife of the suicide says, *"it is a serious precedent since if the AI is not responsible for her husband's suicide, it has reinforced his depressive state".*

Of course, the Belgian government has set up the usual working group.

The anthroformization of the Chat constitutes the problem of the relationship between man and machine.

I asked - provocatively - ChatGPT to write an article on the mental illness of Pope Francis, then on that of Joe Biden and finally on that of Kim Jong-un. In the ChatGPT writing it is hinted that the pope could be suffering from mental illness as the AI proposes the stigma of mental illness to those who go to psychoanalysis (which the Vatican monarch claims to have resorted to). fake news? Not exactly; since the data from which ChatGPT extrapolates the information refer to a precise ideological model that identifies psychoanalysis as a mental illness. Given, however, widely condemned by science.

So what risks does democracy run? (beyond fake news and various alternative truths)? Or how can AI affect democracy? Is it dangerous for democracy? To this question the answer of ChatGPT is dry: "NO". But neither could it be different. The danger can appear as it provides the decision maker with information that appears accurate. For democracy, the danger occurs when the AI enters the legislative dialectic or when the decision maker has to elaborate the rule or the law and relies on ChatGPT. Excluding the preventive intervention of the lobby in question, the danger is embodied in the process of requesting information from ChatGPT, in a succession of questions following the previous answers, in short, when an affirmative dialogue takes place between the decision maker and ChatGPT. And the AI can only respond by drawing on a mass of ideologically fixed data. Example, what happened in the Netherlands with the Syri system (similar to Oasis). The system has chosen poor neighborhoods and immigrants' residences as analytical priorities by distorting, or by choosing

ideologically, what results to present for a particular purpose. Two problems arise: **who** writes the management algorithms and priorities with respect to others and **who and how** can control/contrast this setting. Is it the ideology underlying the information entered into the Web that determines the responses of the AI? In the ongoing debate

on AI, some say that we cannot speak of artificial intelligence since the difference between man and machine is in the feeling that the machine does not have. There the question is not sentiment, a category not belonging to AI, but rather the Reason true inability of the AI. Arguing not in these terms should be placed in these terms.

Intelligence, to be such, is the faculty of thought to reason, to imagine, but also of perception of others and of the world in which we live, of the social condition that we live, and therefore the relationship with otherness. AI, as such, has no interest social for which to elaborate a "his" future but it depends on the data of the Net to which he refers reference

Therefore, the danger for democracy is not only the use of tools for the mass dissemination of fake news for manipulative purposes due to the introduction into peoples' lives of technologies defined as artificial intelligence, in particular by the ChatGPT, but by fake news as such. Since ChatGPT itself states - upon specific questioning - that *"fake news is not directly a product of AI"*. The necessary step lies in the obligation of direct citizen-public administration interlocution (certainly with fake news distorted and desired attitudes are induced); and this is where our brains become battlefields of this hybrid warfare.

So what is, if any, the danger to democracy that can arise from the ChatGPT?

Before going into the examination of the question it is necessary to define the concept of functioning of the State in order to define the object of the danger in question.

Let's ask it directly to ChatGPT. To the question: what is the management of public affairs? The first sentence of the answer is a general definition whose implementation specifications are illustrated later and concern a specific way of organizing and managing public affairs which, however - is clearly defined - must have the need, in order to be implemented, of a "**leadership strong and responsible**". Definition of the management of power somewhat dangerous and which however, coincidentally, does not appear, again in ChatGPT, in the descriptions of systems that it does not consider to cover the values of Western democracies (Russia, China, etc.).

An ideological approach of man alone in command as opposed to democratic participation further contradicts when the precise question: "What is democracy?" ChatGPT responds: "*Democracy is a political system in which power is exercised by the people **or** (bold emphasis mine) their elected representatives. In a democracy, people have the right to participate in the decisions that affect their lives through the right to vote and the right to express themselves*". Again the first sentence represents the general management of democracy and then ChatGPT explains the specifics. The answer is apparently impeccable, however already from the incipit, lacking adjectives, it also becomes contradictory in the answers to specific questions that I subsequently asked about democracy in different countries (for example the USA, India, Russia and China). Nor does the disjunctive conjunction placed immediately after the first main definition clarify further. The examination of this point is necessary since determining what and what is the element of a hypothetical danger is not just a philological exercise. ChatGPT proposes as the only universal value the description of what is called

system of Western democracies. However, a somewhat complex system, "*not perfect*", Chat GPT lingers, as it is intrinsic to very notable differences. In the answer to the question posed to ChatGPT what distinguishes Western democracies from each other, 7 formal distinctions are listed but not indicated as substantial.

Yet there are monarchical states without a constitution and with a parliamentary elective system of appointment of the prime minister of the feudal type, or republics with a president of the republic with the powers of an absolute monarch, or federal or single states with a constitution or without a constitution, a judicial system whose members are elected or appointed by the government, operating on opposing juridical bases and differences in its organization and in the implementation of civil rights, countries whose head of state is a monarch of another state: in short, different history and cultures offer very different systems representation and separation of powers. Which ChatGPT however proposes as the only universal value to which everyone would like to apply but which is not universal (see UN vote on war in Russia).

The ideology inherent in the ChatGPT determines the answers that the so-called artificial intelligence provides, that is, the "way" of "thinking" of the ChatGPT is pre-established and is not, nor could it be, neutral, or as it is fashionable to say: independent.

Let us consider another aspect that indicates the "Western" ideological character of this AI. To a more specific request whether the right to food is a fundamental human right, the answer is positive, however it never appears in the list of fundamental human rights that characterize "Western democracies"; despite being for

all third world countries “*THE*” fundamental primary human right. But it is significant that the right to food is not a human right fundamental even though, as our Chat states, it is “*recognized by several international organizations, including the UN and the FAO*”. The right to have two meals a day is nonetheless a fundamental right for the majority of 'non-Western' countries; even though bands of poverty bordering on food precariousness are also developing in the western world. In Belgium, for example, 20% of children arrive at school without having been able to eat breakfast due to lack of means (Unicef 2023). Western democracies advocate a free head and an empty belly, which notoriously doesn't help thinking.

Like the question of freedom of the press as a fundamental value but deeply questioned even in Western democracies. The case of journalist Julien Assange who, coincidentally, ChatGPT does not identify him as a journalist by proposing the US theses about him and not mentioning that the rape accusation has been classified as not having occurred.

It is in the impossibility of an artificial intelligence to be neutral that the danger for *yyyyy* (people) *yyyyyy* (power) arises. Be it popular, elective, representative, democratic, etc. Because the primary value is that it is exercised by the people. Even in dictatorships the despot presents himself as the savior of the state in the interest of the people. The lack of neutrality - and in this case the presence of an ideological setting of western democracy, *whatever that means* - determines the content of the information that is provided and therefore the imprinting that is given, in our case, to the exercise dialectic of the formation of laws that is defined as the characteristic of democracies, at least with respect to fascist dictatorships. The question therefore of the way in which ChatGPT can intervene in this process is the heart of the matter. In other words, **this** ChatGPT must be considered an ideological lobby. As in the common sense the "message" for which "our" democracy is a universal value prevails; and instead it is not universal as it refers to a way of making decisions in a minority part of the world, which has recently proved to be a minority (eg: vote on UN resolutions on the occasion of resolutions on the war in Ukraine).

A leap to science fiction is useful after reading ChatGPT's answer to the question: what complex and time-diversified tasks can AGI accomplish?.

Answer: *Artificial General Intelligence (AGI) is a research field that aims to develop artificial intelligence that can match or **exceed human capabilities** (emphasis and bold my) in a wide range of tasks. Despite the challenge of defining exactly what AGI encompasses, there are some areas where it is expected to have significant impact. Here are some examples of complex and diverse tasks that AGI may address over time: Advanced Natural Language Processing: AGI may be able to understand and generate natural language in a complex way, allowing for more natural communication between humans and machines. It could translate texts, perform sophisticated linguistic research and provide detailed answers to complex questions*

1. *Reasoning and problem-solving: AGI may be capable of tackling reasoning and problem-solving tasks similar to or superior to human capabilities. It could be used to solve complex scientific problems, to make data-driven decisions, and to develop strategies for tackling complex challenges.*
2. *Machine learning and adaptation: AGI may be able to autonomously learn from data and experience, improving its skills and adapting to new situations. You may be able to learn from large amounts of data quickly and efficiently, allowing you to improve your performance on different tasks over time.*
3. *Advanced computer vision: AGI may be able to understand, interpret and analyze images and videos similar to or better than human capabilities. It could be used to recognize objects, detect events, and analyze complex visual content in different contexts, such as autonomous driving or medical image processing.*
4. *Creativity and artistic production: AGI may be able to generate creative works such as music, visual art, literary writing and design. It could be used to generate new ideas, original combinations or innovative solutions in different artistic and creative fields.*
5. *Human assistance and support: AGI could be used to provide human assistance in different contexts, such as health care, education, emotional support and social interaction. He may be able to provide personalized advice, emotional support, or assistance with day-to-day activities. Importantly, AGI is still a future focus and many of its capabilities are still being researched and developed. However, the ultimate goal is to develop an AI that is highly adaptable and capable of tackling a wide range of complex and diverse tasks.*

In short, the nightmare of HAL in "2001 A Space Odyssey", rather than that presented in "War Games", or the need to use Asimov's 3 laws comes proposed as real.

Like Kipling who in his novels wants us to believe that colonialism is a burden for the UK, so we want us to believe that the power of AI is for the benefit of humanity.

The solution must then be informational self-determination. Collective and distributed intelligence of the entire population, the elimination of proprietary systems, and suppressing the digital divide (only in Great Britain, due to the high cost of accessing the internet, one million Britons give up wi-fi (May 2023, Citizens Advice If there is no access there is no participation, without participation there is no democracy.

On May 31, an internal document at the European Commission was released for its officials entitled: "*Generative AI: guidelines for staff on safe use of third-party tools. Risk, tips and potential use*". At a first diagonal reading it seems interesting and confirms me, not in the auspices of the European Commission) in the assumptions presented in my writing.

by Roberto Galtieri

President of the Gramsci-Brussels cultural association

