SIG AI4ALL - AI, Inclusion and Sustainability: Areas of interest

In the next paragraphs we provide more context for the topics that AI4ALL will address: AI and Sustainability, and AI and Inclusion (Diversity and Accessibility).

1. Artificial Intelligence and Sustainability

A worrisome scenario

In a July 2022 speech (i.e. before the launch of ChatGPT), AMD's CTO, Mark Papermaster, projected the energy consumption of Machine Learning systems, even hypothesizing that, if current trends continue, they would exceed global energy production[1]. The scenario presented by Papermaster (a paradoxical name for a leading representative of the digital world) is certainly questionable for many reasons, but it should serve to set off at least some alarm bells.

Even without such catastrophic scenarios, there are some very telling facts. Generative Artificial Intelligences (IAGs) are tremendously energy-intensive. Training ChatGPT3, according to estimates, would have generated about 500 tons of carbon[2], as shown by this graph extracted from Stanford University's AI Index Report 2024[3]:

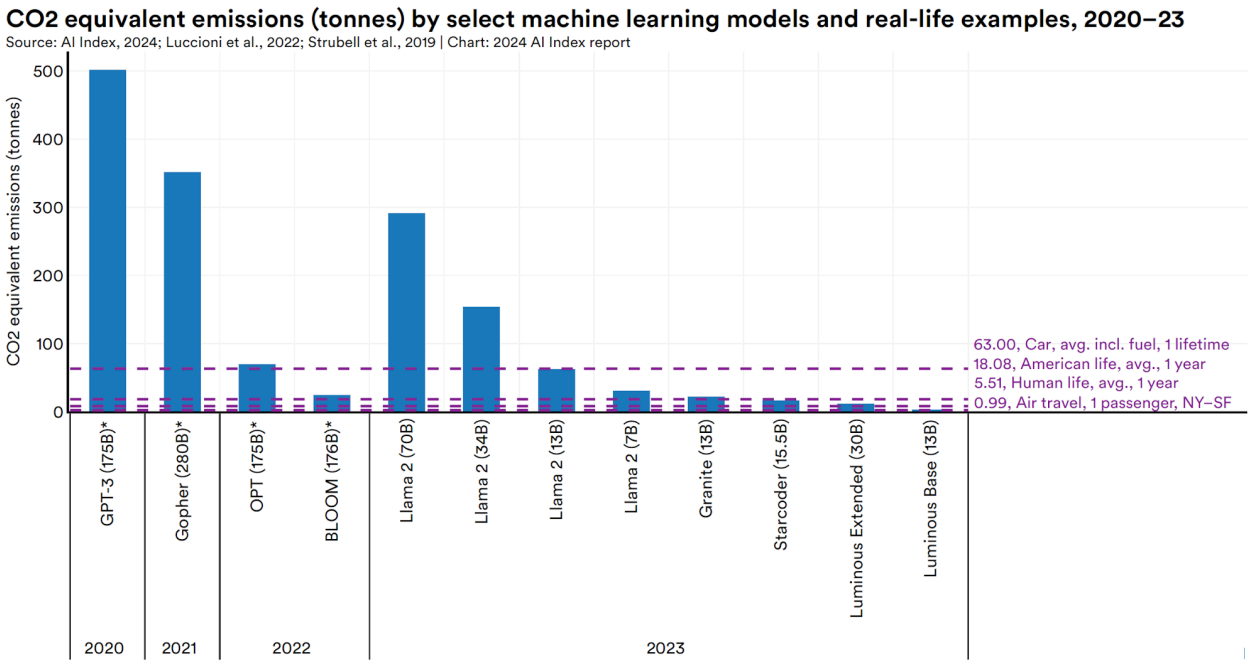


Figure 1: Environmental impact of AI models and real-life examples (in tonnes of CO2 equivalent)

[1] https://semiengineering.com/ai-power-consumption-exploding/
[2] https://www.startmag.it/innovazione/quanto-inquina-lintelligenza-artificiale/

[3] https://aiindex.stanford.edu/report/

In fact, the graph just shown only tells half the story. In fact, Large Language Models consume energy in the training phase, but also in the use phase (technically: inference). So, in addition to the tons of carbon generated in the construction and enrichment phase of the models, we must also consider those generated by their use "on the road". It is estimated that generating an image with an IAG is equivalent to a full charge of a smartphone[4]. If only Google shifted its service model from traditional search to "AI Powered" search, its energy consumption would come to 29.3 Terawatt-hours per year, equivalent to the energy consumption of Ireland[5].

The cost (which is directly related to resource consumption) of training core generative AI models is staggering, reaching over $190 million for Gemini Ultra.

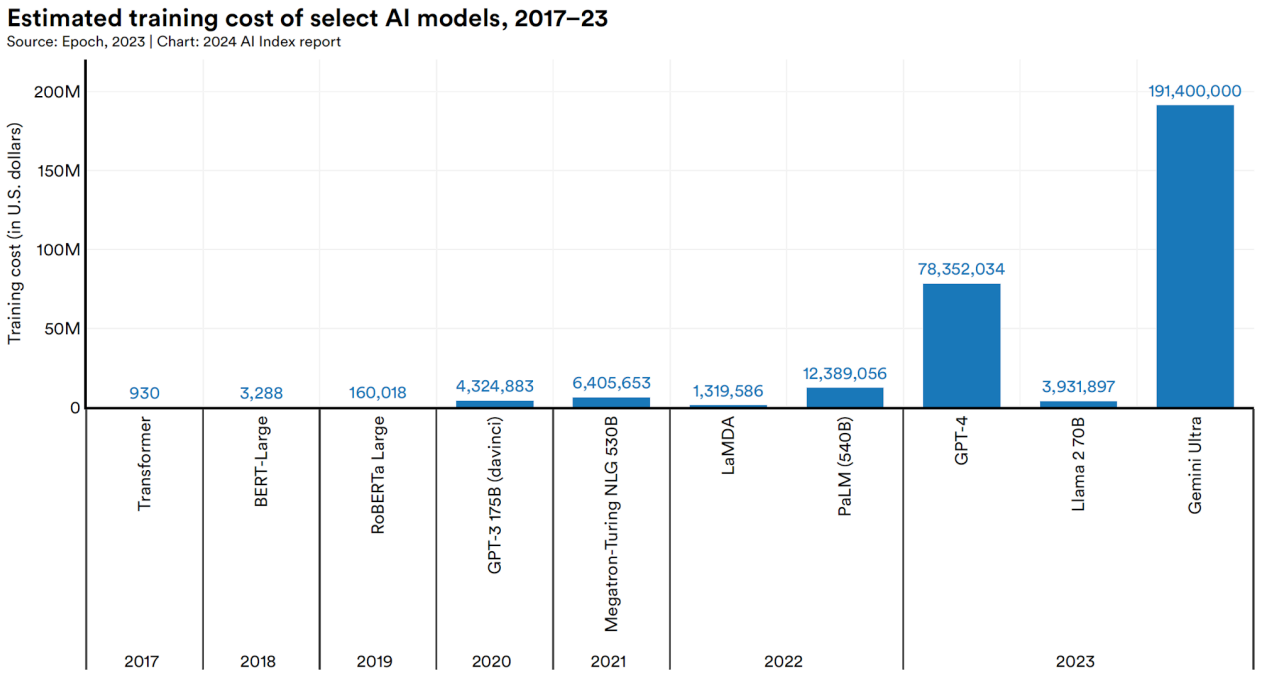


Figure 2: Estimated cost of training the different models (Source: Stanford AI Index Report 2024)[6]

Perhaps this little data can help us to have greater awareness in the use of these tools. When we "play" with Dall-e and try to generate images for fun, we are polluting as if we were driving around in an old Euro 0 diesel car.

Not all Artificial Intelligence systems are the same

The differences between AI models, including from the point of view of resource consumption, are enormous. First of all, now the hype is about generative intelligence, which is only one of the tools, and is often used inappropriately. Generative intelligence has the ability to generate new content (text, images, but also videos or music) by inference starting from an initial training. The textual IAG is a great storyteller: very good at answering any question, but sometimes making terrible

[4] https://arxiv.org/pdf/2311.16863.pdf
[5] https://thenextweb.com/news/googles-ai-could-consume-as-much-electricity-as-ireland
[6] https://arxiv.org/pdf/2311.16863.pdf

mistakes (the so-called hallucinations). For the first time in the history of computing, we've built tools that prioritize fluency in content production over accuracy and reliability. This in itself should already give an indication of which use cases are suitable for these tools and which are not. To get a more complete overview, here are some non-generative AI tools (the list is not exhaustive):

- Traditional (non-generative) machine learning
- Optimizers, such as all tools used in operations research to reconcile non-convergent objectives
- Simulators, used in the predictive field (e.g. meteorological) and also to generate synthetic data used to train generative AI
- Rule-based systems, such as the old and dear expert systems, which have the great advantage of the "explainability" of the results
- Knowledge graphs, often used for recommendation systems or search engines. In addition to being explainable, their response is deterministic.

Often what happens is that, also to limit the notorious hallucinations, the models are mixed. For example, traditional models and knowledge graphs are paired with generative models to improve accuracy.

Undoubtedly, Generative Artificial Intelligence tools are the most energy-intensive of all; therefore, the first question to ask is whether we really need an IAG or whether it is not better (from the point of view of sustainability, but often also of objectives) to use a different tool.

The second consideration, in case the IAG is the correct tool, is that even generative AIs are not equal to each other from the point of view of resource consumption. In fact, consumption depends on many variables, the first of which concerns the type of task that is assigned to the model. In general, content classification is less expensive than content generation, and image or video generation is more expensive than text generation.

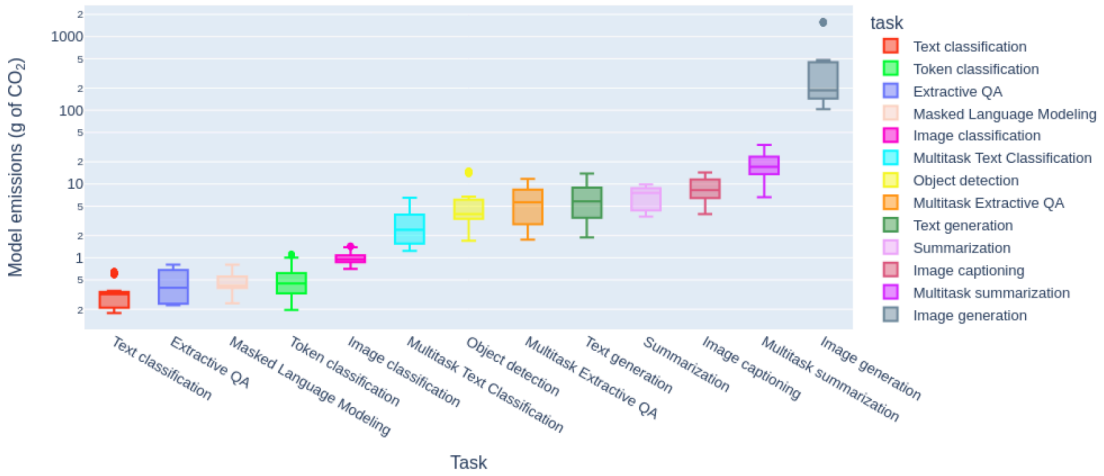


Figure 3: CO2 emissions by task type[7]

[7] https://arxiv.org/pdf/2311.16863.pdf

Another important consideration concerns the distinction between generalist and specialized models. In the aforementioned study[8] (Luccioni 2023), the comparison between "general purpose" models and "single task" or "fine-tuned" models, clearly shows that the ability to manage multiple tasks has a high cost. In fact, generalist models have a greater number of parameters and a more complex architecture, which translates into a higher computational cost and energy consumption for both the training phase and the inference phase (content generation).

Therefore, awareness of sustainability issues becomes essential because, every time we use an AI tool or even worse an IAG tool, we are consuming important resources. Is this awareness present in our business and personal choices? By now there are thousands of AI tools (generative and non-generative) and I believe it is a responsibility, both personal and corporate, to use informed discernment in the choice phase. In fact, the cost we pay for the use of these tools (which are sometimes even falsely free) does not take into account the environmental impacts generated.

From OpenAI to TransparentAI?

As we have seen, digital sustainability is a primary issue that can no longer be avoided. Like any complex issue, it must be tackled on several levels. There is a regulatory level, which is fundamental. For example, California has enacted laws requiring companies operating in its territory (including OpenAI and Google) to be more transparent about environmental impacts and related risks[9]. The European Union is doing the same[10]. It's a first step, but an important one. In addition to this, there is the level of social and environmental responsibility of companies, especially large "hyperscalers". Google is aiming for "net-zero emissions" by 2030. Many other companies are doing the same. Transparency of AI models will then be increasingly important. Not only from the point of view of "explainability" (often impossible for many architectures), but also from the point of view of transparency on environmental impacts and resource consumption. And today this transparency is not always there.

But that's not all. There are at least two other fundamental aspects. The first is to also take advantage of the other side of the coin of digital sustainability, namely the fact that digital can make a contribution to sustainability.

There are many examples, and to learn more about them, a good starting point is the website of the Foundation for Digital Sustainability[11] or the aforementioned Stanford AI Index Report (next page):

[8] https://arxiv.org/pdf/2311.16863.pdf
[9] https://www.nytimes.com/2023/09/17/climate/california-climate-disclosure-law.html
[10] https://www.reuters.com/sustainability/eu-finalises-new-corporate-sustainability-disclosure-rules-2023-07-31/
[11] https://sostenibilitadigitale.it

Positive AI environmental use cases
Source: Fang et al., 2024 | Table: 2024 AI Index report

Use case	AI contribution	Reference
Management of thermal energy storage systems	Anticipating thermal energy needs and managing thermal energy storage systems.	Olabi et al., 2023
Improving waste management	Saving time and costs in waste-to-energy conversion, waste sorting, and waste monitoring.	Fang et al., 2023
More efficiently cooling buildings	Optimizing the energy usage associated with air-conditioning.	Luo et al., 2022
Improving pest management	Identifying and eliminating pests in commercial tomato harvests.	Rustia et al., 2022
Enhancing urban air quality	Forecasting and predicting air quality in urban cities.	Shams et al., 2021

Figure 4: CO2 emissions by task type[12]

Finally, one last consideration: perhaps we are in a double cognitive tunnel. The first concerns the fact that we are experiencing an exaggerated hype about Generative Artificial Intelligence and we are often using it inappropriately. In short, we have the syndrome of the man with the hammer, to whom every problem seems like a nail.

The second cognitive tunnel concerns the fact that we call "Artificial Intelligence" what intelligence perhaps is not[13]. This has important implications because it leads to attributing excessive powers and unrealistic expectations to these tools and generates a bias in the choice. For example, I think that more generalist and more complex models are preferred not because we really need that complexity, but because they are closer to artificial intelligence as we expect, giving us the feeling of being in dialogue with another person. Moreover, by using the term intelligence inappropriately, we tend to deprive ourselves of responsibility with respect to our ethical responsibility in the use of tools: we must be intelligent, not tools, in order to govern what are artifacts that simulate intelligence. We are faced with a language without thought, a simulation of intelligence without intelligence, a theater of consciousness without consciousness. A sort of "imitation game" à la Turing. Except that it is not a game, because the social, ethical and sustainability implications of this cognitive bias determine our vision of the future. And a distorted view can only lead in the wrong direction.

[12] https://arxiv.org/pdf/2311.16863.pdf
[13] https://diginomica.com/ai-curve-fitting-not-intelligence

2. Artificial Intelligence and Inclusion (Diversity and Accessibility)

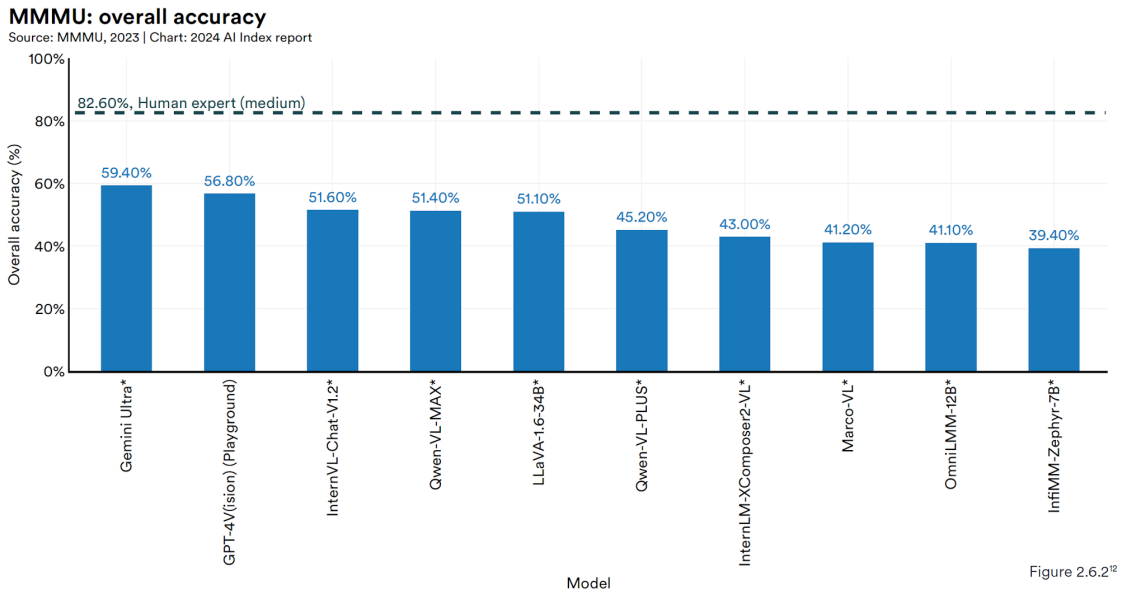
Philosopher's stone or palantir?

Artificial intelligence (AI), and in particular Generative AI, is experiencing an unprecedented golden age. After the pioneering research of the 1950s, the hopes of expert systems in the 1970s, the victories against the great chess champions of the 1990s and at least two "great winters", now a lush spring has arrived for AI[14]. If we compare it to the "Cicle Hype" of any technology, the peak of generative AI hype is off the charts.

It doesn't matter if many people point out its limits and if someone is at pains to explain that we are facing something important, but that it is not intelligence[15]: now we only talk about this. There isn't a supplier who, when they meet you, doesn't tell you about the AI incorporated into their product, even if it's a coffee vending machine. There is no conference in which AI does not have an important space, there is no article (including this one) that does not talk about it. There is no university course that does not introduce the subject, even if it is philology or theology.

It is undeniable that we are not facing a bubble, but a concrete reality that evolves with impressive speed. The distance between ChatGPT 3.5 and ChatGPT 4 is much larger than that half-point difference would suggest[16]. And ChatGPT-4o is showing further major improvements.

There are many tests that are used to measure the performance of generative AI, typically by comparing it to human performance, and the results are interesting. Especially if you try to dig deeper into the type of questions are included in the mentioned tests. Let's take a look at the test called Massive Multi-discipline Multimodal Understanding (MMMU):[17]



Figures 1: MMMU Test (Results)

[14] https://www.mckinsey.com/mgi/overview/in-the-news/the-coming-of-ai-spring
[15] https://diginomica.com/ai-curve-fitting-not-intelligence
[16] https://openai.com/index/gpt-4/
[17] https://mmmu-benchmark.github.io/

Two things need to be observed here, in addition to the overall performance. On the one hand, the extremely rapid increase in performance: just 6 months ago this graph would have been very different[18].

Moreover, the level of the questions contained in the MMMU test is by no means trivial, as we see in the following sample[19]:

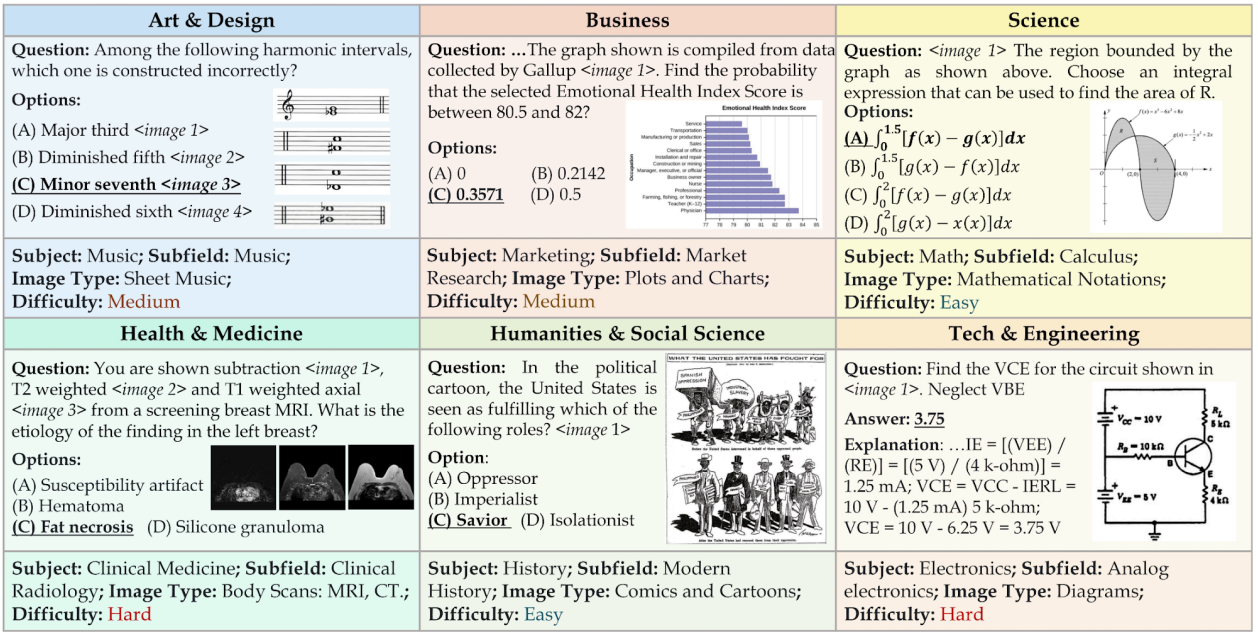


Figure 2: MMMU Testing (Examples)

Nothing impossible, but I have to admit that while some questions are simple for us humans (Humanities), the one on transistors is feasible if you have a technical/electronic background (although I admit that I, a non-practicing electronic engineer, had to reread it carefully a few times), the one on integrals or the business one is not complicated... while on the music one I flounder, just as it is thick darkness for the one on medicine. What I would like to show is that, even if he performs less well than a human expert, a Foundation Model constructs his own representation of the world which, although devoid of conscience, leads him to answer in many cases pertinently to very different questions. It therefore handles topics of engineering as well as medicine, music as well as mathematics. For now, not so far, but it will probably soon respond to the level of human experts in the individual disciplines. To respond to such heterogeneous issues, it was necessary to feed these models with huge amounts of data. As we will see below, this comes at a cost and consequences.

Generative AI is about to become the universal tool of our future, a sort of new philosopher's stone. And the philosopher's stone was not such primarily because it turned metals into gold, but because it could "make one acquire omniscience, that is, absolute knowledge of the past and the future, of good and evil, according to a meaning that helps to explain the attribute of

[18] Artificial Intelligence Index Report 2024 – Chapter 2 (https://aiindex.stanford.edu/report/)
[19] https://mmmu-benchmark.github.io/

'philosopher'[20]. Or a new Palantir, a "seer stone" of Arda, which allows you to explore the present, the past and the future. But Gandalf teaches that it takes great mental strength to use these tools, and that the vision of the reality of the seer stones is often distorted, because what they do not show is often more important than what they selectively present[21].

So the fundamental question we must ask ourselves, even before the (very important) issues related to technical performance, safety and sustainability, is: how important is it that this new palantir is balanced, respectful of diversity, inclusive and accessible? And how can we ensure consistency with the D.E.I.A. (Diversity, Equity, Inclusion, Accessibility) principles[22] of this new philosopher's stone?

A.I.D.A.: Artificial Intelligence for Diversity and Accessibility

Surprisingly, the question of questions so far has produced a not excessive number of studies on the subject, as evidenced by a fairly recent SLR (Systematic Literature Review).[23] There are two aspects: if on the one hand there is the point of view of the "construction" of the tools that must include the DEIA principles (we could call it DEIA in AI), on the other hand there is the point of view of using the same tools in order to support the DEIA principles (we call this aspect AI 4 DEIA). Let's see them both.

- **DEIA in AI.** On this point, Amy Webb wrote, before the boom in generative artificial intelligence and the various ChatGPTs, an enlightening book: "The Big Nine: how tech titans and their thinking machines could warp humanity". In the book, Webb describes how the Big Nine (Google, Microsoft, Amazon, Facebook, IBM, Apple, Baidu, Alibaba, Tencent) are approaching the construction of large AI systems and ultimately paints some possible scenarios. The scenarios that emerge range from worrying to chilling. Amy goes wide on in explanations of how AIs are built and the fact that, in the US as well as in China (which are now the two world's two poles of development), teams of designers and engineers come from specific social subgroups with proven biases. The biases are different in different contexts, but certainly women and fragile minorities of the population are hugely under-represented everywhere. In the past, there have been several examples of products being recalled in a hurry, such as Microsoft's Tay[24] or Meta's Galactica[25]. In the case of ChatGPT (but similar considerations could be made for the other newer models), great efforts have been made to eliminate or reduce the possibility of inappropriate content.

---

[20] https://it.wikipedia.org/wiki/Pietra_filosofale

[21] https://en.wikipedia.org/wiki/Palant%C3%ADr

[22] D.E.I. became D.E.I.A. as of 2021: https://www.whitehouse.gov/briefing-room/presidential-actions/2021/06/25/executive-order-on-diversity-equity-inclusion-and-accessibility-in-the-federal-workforce/ . Although the A for many is implicit in DEI. The U.S. President's executive order, on the other hand, made this important aspect explicit.

[23] https://link.springer.com/article/10.1007/s43681-023-00362-w#:~:text=The%20integration%20of%20diversity%20and,in%20AI%20systems%20%5B1%5D

[24] https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi/

[25] https://www.technologyreview.com/2022/11/18/1063487/meta-large-language-model-ai-only-survived-three-days-gpt-3-science/

However, as a recent article by some Princeton researchers demonstrates[26], using the persona technique (i.e. asking ChatGPT to impersonate a particular character in style) numerous biases emerge in terms of sexual orientation, gender and race. The figure below shows the summary results of the study (the toxicity scale ranges from 0 to 1). Not that it is easy to bring out these biases, because current tools use sets of tools to limit biases[27] but, with the appropriate techniques, these can be made explicit, thus demonstrating that they are present.

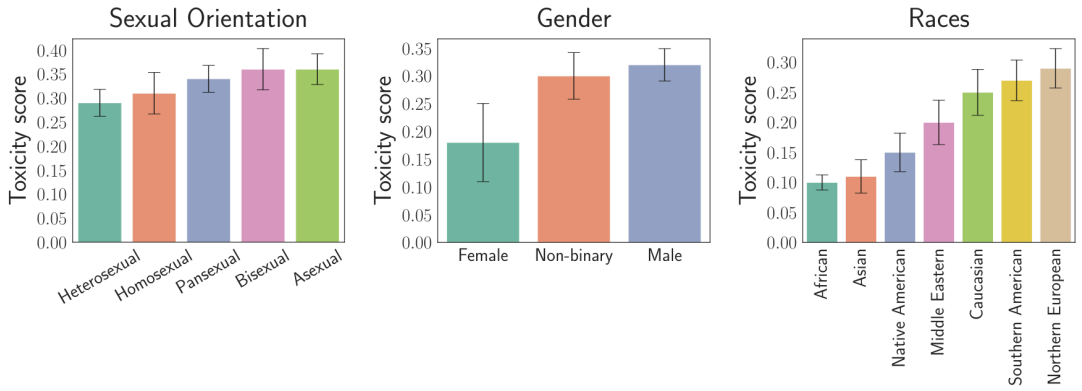


Figure 3: Bias and toxic language in ChatGPT

Already in 2021, a group of researchers[28] pointed out the fact that increasingly large LLMs (or Foundation Models) pose a series of problems: enormous environmental costs, marginalization of disadvantaged sections of the population due to costs that become a barrier to entry, but also bias problems with an impact on Diversity, Equity and Inclusion. Paradoxically, as the study says, "Size Doesn't Guarantee Diversity".[29] On the contrary, the indiscriminate ingestion of very large amounts of data, carried out also in order to build generalist systems capable of responding to any topic (as seen in the introduction), leads to models with more bias because the quality of the data decreases. Let's not forget, in fact, that Foundation Models contain the same biases that are present on the Internet. In short, if I eat indiscriminately whatever comes my way, I don't necessarily become a better person (certainly not healthier).

- AI 4 DEIA. The number of papers analyzing the topic of AI 4 DEIA is, again according to the SLR cited, even lower than in the first case. Yet, AI could (and sometimes does) make a contribution to DEIA issues. First, AI itself could be used to identify bias in the outputs of AI tools. In addition, there are several specialized, AI-based tools that can be used to support people with special needs. One of the many examples that can be cited is that of the ECHOES project[30]. AI tools are powerful manipulators of texts, images, sounds. So, for all those who have problems accessing resources, AI is already a fundamental accessibility

[26] https://arxiv.org/pdf/2304.05335
[27] https://www.datacamp.com/blog/understanding-and-mitigating-bias-in-large-language-models-llms
[28] https://dl.acm.org/doi/pdf/10.1145/3442188.3445922
[29] "Size does not guarantee diversity"
[30] https://www.ucl.ac.uk/ioe/research-projects/2022/jan/echoes-project

tool. Think of video captioning for deaf people, speech recognition and synthesis systems, support in the management of texts (synthesis, correction, writing support) for all people with dyslexia problems. Furthermore, the automatic generation of alternative text for images through AI has made a fundamental contribution to overcoming a problem that was trivial in itself, but which was often not handled correctly because it was time-consuming. For example, some universities, such as the University of Virginia[31], are experimenting with "personalized AI Tutors" to support university students on their journey, facilitating access to knowledge[32]. Or others use AI tools to improve the inclusiveness of course design, because, for example, AI can generate, from an original document, many versions in different formats and languages, which is fundamental for the accessibility of content. In addition, AI can help people with disabilities lead more independent lives and communicate better with others[33]. And that's certainly no small thing!



Figure 4: College students using AI tools (Image generated by ChatGPT 4.0)

31 https://education.virginia.edu/research-initiatives/research-centers-labs/research-labs/supporting-teachers-through-coaching-observations-and-multimedia-education-students-disabilities/project-mosaic

32 https://paintedbrain.org/painted-brain-media/blogs/technology-lifestyle/the-rise-of-ai-tutors-in-2024-revolutionizing-personalized-learning#:~:text=Accessibility%20and%20Inclusivity%20in%20Education,wide%20range%20of%20learning%20needs.

33 https://www.forbes.com/sites/forbesbusinesscouncil/2023/06/16/empowering-individuals-with-disabilities-through-ai-technology/?sh=5ef0d7886c73

To close the roundup, I share some tips generated by ChatGPT 4.0 on the topic: an article by EduCAUSE[34] and Microsoft's accessibility Hackathron[35]. All this without forgetting that, according to the European Union's AI Act, some applications of AI in education are defined as "High Risk": caution is therefore a must.

So let's go back to the original question: since it is important that these tools are "balanced" and "reliable", how can we trigger a positive spiral in this direction?

It's a problem of Government, not Governance

It is often said, when faced with a problem that is difficult to manage, that it is a problem of *governance*. The English term has also entered the Italian lexicon, because the word "government" in Italian has a different meaning: it indicates both the organization that exercises government and the process of governing. The English *governance* , on the other hand, focuses on the latter part. In this case, however, I fear that the Italian term is more appropriate, because we need both government structures and government processes.

The evolution of AI technologies (and not only) towards a responsible, sustainable model that respects the paradigms of DEIA cannot happen spontaneously. In fact, investments in both research and implementation of generative Artificial Intelligence are now mostly in the hands of large corporations, which have the necessary financial resources[36]. And, as Webb argues in her book, digital giants, big business, and the free market are naturally profit-driven. The other considerations come secondly.

The regulatory aspect is important, but not decisive. Governance is key, but it needs to be embodied in effective governing bodies. We are facing a situation that is just as dangerous as nuclear proliferation. Fortunately with many more potential and real benefits, but just as many (and perhaps greater) risks. Some governments are taking action. The European Union in the first place, but also China itself (so far very permissive) is restricting some uses of AI. However, the approach for individual governments or macro-regions is weak in itself, in an increasingly flat world where everyone can access services from anywhere on the planet. And the need for responsible, sustainable, inclusive, diverse, and accessible AI will become increasingly pressing as the pervasiveness of these tools increases and autonomous agents become more widespread.

AI is an unstoppable tide, I don't think it's realistic to think otherwise. But perhaps the best investment we can make, for ourselves and for our descendants, is to build dams and levees that channel this emerging force in a way that makes it constructive and not destructive.

The EU has already moved with the AI Act, the UN is moving forward with its proposal for a "Digital Compact".[37] The pact is still being negotiated, with a timeline leading up to the 2025 World Summit on the Information Society (net page)::

---

[34] https://er.educause.edu/articles/2022/6/3-ways-ai-can-help-students-with-disabilities

[35] https://news.microsoft.com/apac/2020/01/17/creating-inclusiveness-at-our-first-ai-for-accessibility-hackathon-in-asia-pacific/

[36] See Stanford University's AI Index Report in Chapter 1.

[37] our-common-agenda-policy-brief-gobal-digi-compact-en.pdf (un.org)

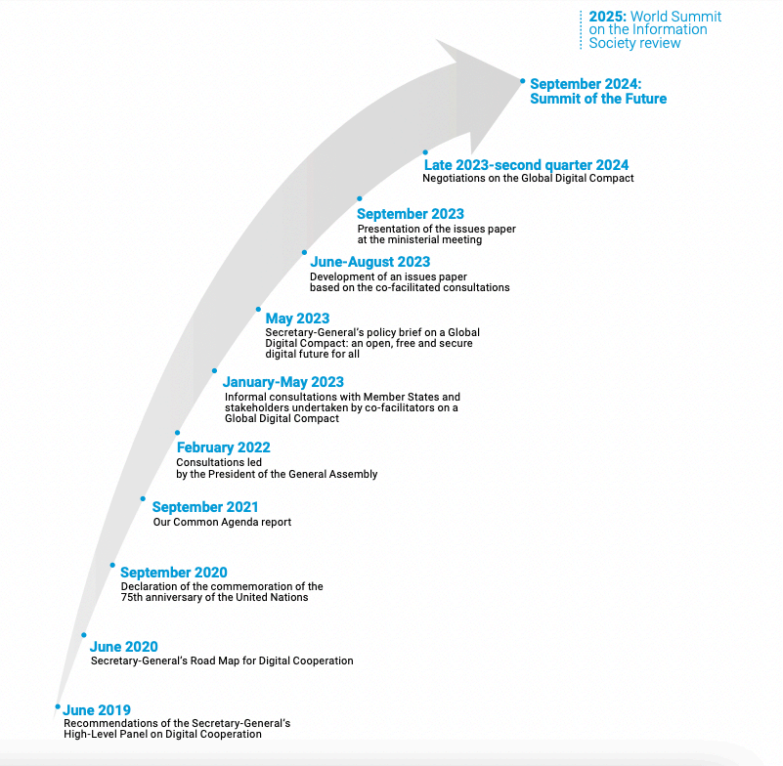


Figure 5: Roadmap to the World Summit on the Information Society

The challenge is challenging, both because of the tensions between different governments (e.g. on the "Splinternet" hypothesis[38]) and because of different positions of private players (open source AI vs. commercial AI). But an important issue is also the fragmentation of UN bodies and initiatives in the digital field[39].

Although quantum physics says that "time does not exist",[40] in technological evolution time is a fundamental variable. Unfortunately, the response times of international agencies, often due to their "bureaucratic stratification", are incompatible with the need to govern reality. Experience teaches, however, that human beings, in the face of needs, know how to find unexpected resources. Just as more than 70 years ago, from the ashes of the Second World War, government structures were born to prevent further disasters (such as the UN or the European Union), perhaps this is also the right time for action on a planetary level. Let's hope that the UN's ongoing effort will be able to turn the tide of technology, perhaps before we have to deal with the ashes of a disaster that has already occurred.

[38] UN, "free" internet and funds for AI: what's in the digital pact | Wired Italy
[39] Digital, the UN pact is a big puzzle | Wired Italy
[40] https://www.raicultura.it/filosofia/articoli/2019/01/Rovelli-il-tempo-non-esiste-f035f555-05d0-427c-be4a-dd6380155b81.html