

Grant G. Buffett

Terranova Scientific, C. Serralet, 70, Masquefa (Barcelona), 08783 Spain

## GENERATIVE ARTIFICIAL INTELLIGENCE IS CHANGING THE LANDSCAPE OF RESEARCH AND EDUCATION

### Introduction

During my undergrad in Earth Science, now some 30 years ago, I enrolled in one of the most memorable courses of my life: Sociology (Technology & Society). Since that time, I have remained fascinated with the intersection of science and society. So, over the past few months, it has become abundantly clear to me that we are on the cusp of another step change in how we use technology. I am referring to the advent of Generative Artificial Intelligence (Gen AI).

It was early December 2022 and I had just signed up for ChatGPT. Even after having known of its potential for several years, I was stunned by how competently it could write. After using it for one year and having delved more deeply into how it is implemented, I see this new technology as an invaluable tool for researchers. However, as the technology is advancing quickly, we must soon implement responsible regulation to maximize its benefits, while successfully navigating its inherent pitfalls.

I use GPT to ask questions or probe ideas, knowing that it isn't an oracle (yet) and that it is my responsibility to read its responses with a critical eye. I have found that GPT excels at some topics while failing miserably at others, getting even some simple information wrong, or even inventing explanations. Sometimes you just need to tell it that it is wrong, and it will accept that criticism and agree with you. But even where it gets things wrong, its grammar is nearly perfect. For that reason alone, it can be an especially valuable writing tool for researchers.

Although I have spent most of my professional career as a seismologist, over the past six years I have also freelanced as a copy editor. In this respect, I first saw GPT as an immediate threat to my editing business. However, considering its virtual inevitability going forward, it became quickly clear that I would need to add GPT to my editing toolbox—in much the same way that we have all adopted web search, Google Translate, word processors, spell-check, or calculators. GPT can be your personal intern or graduate student, taking care of much of the tedious work for you. For example, you might task it with fetching information, compiling data, helping to brainstorm ideas, helping prepare for your upcoming debate, playing the role of a peer reviewer, or analyzing datasets for possible correlations.

Since the normalization of mobile devices, we are all first-generation cyborgs, augmented with the world's information in our pockets. GPT can be the tool to parse and make sense of that information. But can we use it to truly create knowledge and understanding? In the following, I will expand a little more on Gen AI (of which GPT technology is a part), speak about its positive and negative societal implications, and leave you with some ways in which it can be useful to research and education.

### The natural language processing revolution

GPT stands for “Generative Pre-trained Transformer.” *Generative* because it has the ability to generate novel information (both right and wrong), *Pre-trained* because it doesn't need to search the internet for its answers (although the normal way is to access it via the internet), and *transformer* refers to the underlying architecture used in GPT. Transformers are deep learning models that excel at tasks involving sequential rule-based datasets, for which language happens to be a good example. The transformer architecture employs a self-attention mechanism, allowing it to capture relationships between words in a sentence and to efficiently “learn” dependencies. You can think of it as a very sophisticated version of auto-complete.

GPT-based large language models (LLMs) are now poised to become so inexpensive as to simply become part of the Internet of things. In terms of file size, they are already optimized enough to be run on a desktop computer. GPT will soon be as normal as a Google search. But, while all that is happening, GPT, and LLMs in general, will be very disruptive to our modern information space.

Automatically generated and unverified news articles have already started to appear and will do so more frequently. This will further erode society's trust in our established scientific and our political institutions. Consider for a moment communication about health or new and emerging treatments for diseases. In practice, how will we know if online content is AI-generated or is just click-bait? There are online GPT output detectors, but most people will not bother to regularly verify information in that way. Moreover, such detectors are imperfect and will constantly need to be updated to deal with the evolving LLMs.

In a recent podcast episode of Making Sense<sup>[1]</sup>, philosopher and

neuroscientist Sam Harris interviews Nina Schick <sup>[2]</sup>, a Gen AI expert, who has advised many tech corporations, governments, and the United Nations. Considering the exponential growth of information, she predicts that over 90% of online content will be AI-generated by as early as 2025. This implies the potential for a sharp rise in misinformation and disinformation. The most straightforward solution to its spread is a technical one—simply create the detection algorithm and use it to embed authentic files with an encrypted, permanent digital certificate. We already do this for many official documents. The caveat to this solution is that, while straightforward, it is technically tedious to implement en masse. In addition, at present, detection algorithms only give a percentage likelihood as to whether fake content is a false positive. Moreover, we are quickly entering into the realm of multi-modal foundational models. So, it's not just text documents we have to be concerned about, but also text combined with video, digital art, and voice synthesis. Apple's latest iPhone has built-in voice synthesis software that just needs to sample your voice for a few minutes to be capable of convincingly emulating it. Will we soon have to anticipate phishing calls in the voices of our loved ones, asking for money or help!? Given the growing volume of data, it seems likely that our information space will soon not only be generated by, but must also be *verified* by other AIs. If so, with which institutions do we place our trust?

Optimistically speaking, if we manage it well, AI could be an immense benefit to society. However, in practice, the new capacity to create content at scale suggests that digital media will increasingly be at risk of losing credibility. Moreover, where newly generated AI content is financially incentivized, we run the risk that it would skew toward simply whatever gets the most user engagement (easily measured by clicks or scrolling/pause speed). This implies that AI creations will bias toward emotional content (including but not limited to fear, rage, hate, revenge, sexual desires, etc.). To maintain a civil public discourse, we will soon need responsible government regulation.

For most, our daily information diet is already heavily biased toward digital content over print. If we extrapolate this trend, it becomes clear that the first big change we need to make across societies is to adapt our school system from a young age to prioritize critical thinking over the brute memorization of facts. Children need to be taught that it is fundamental to ask basic questions and be naturally skeptical about new claims. For example, what is a source of this claim? How credible is it? Who benefits from a claim being true? Secondly, we need to prioritize statistical thinking in schools, because, to analyze the veracity of a claim, we must be able to deduce its likelihood. We need more first principles thinking. Considering the growing flood of available information, if we don't make this step-change to our educational systems soon, we will remain susceptible to all types of deception.

Gen AI models and algorithms are the worst that they will ever be. While artificial intelligences probably haven't yet passed the Turing Test (the ability to convince a human that they are *not* an AI), they do come close. And, if you're not fooled by an AI, it might be paradoxically because there is an obvious giveaway: the unhuman rate at which it can generate coherent information. There is already an app that is task-oriented and can

automate decisions for you, called AutoGPT <sup>[3]</sup>. Barring some economic or energy crisis, Gen AI technologies will continue to rapidly improve.

### Applications and limitations of GPT for researchers

Let's focus on one aspect of Gen AI: Natural Language Processing (NLP), of which GPT is one implementation. As researchers, we need to review the literature, manage and analyze datasets, and communicate results in a clear and concise manner. Here are some ways that GPT can help.

In general, I would recommend that you prepare GPT for your input by giving it some information about what you want to accomplish. Explain a little about yourself, and what are your goals. GPT really excels when you place clear constraints on its output (i.e., number of paragraphs, words, or characters, etc.). Consider asking it to format the text in a particular style or publishing standard. Ask it to write about a topic, but designed for a certain audience (children, colleagues, the general public, etc.). You can also paste text from other sources and ask it to analyze, modify, translate, or condense it. In the latest version of ChatGPT you can upload PDF files of a published article, for example, and then discuss its results together. You can ask it to role-play, taking the perspective of another person, profession, or a peer reviewer (while I don't expect AI to replace the peer review process soon, it's a good dress rehearsal before submitting your article to a journal). The more accurate information you give it, the more suitable will be the output. Most interestingly, GPT can even code or debug for you. Ask it for some Python or C++ script for a specified purpose (or create Excel spreadsheets or macros), and you'll have yourself a working template, which you can then modify to suit your needs.

For my part, I mainly find it useful to write outlines on a particular topic (something that prevents many researchers from getting started). I employ it to analyze text for some hidden detail or trend, to simplify jargon-laden documents in the language of your choice, or to suggest alternative phrasings.

GPT (or more sophisticated apps like AutoGPT) could help researchers to conduct literature searches by analyzing keywords and automating the retrieval of relevant scientific papers from online repositories. It could then help to summarize and synthesize key information from multiple sources, making the review process more efficient (if somewhat black box). Of course, the ultimate decision to use the information must be up to you. The dilemma for many researchers will not be due to a lack of information, but rather saturation by it. GPT will likely be the cause of (and solution to) such dilemmas. You could employ it to generate a wealth of information on a certain topic (brainstorming ideas), to synthesize and reduce that information to a given focus, and to organize those thoughts and ideas by generating outlines (introduction, methodology, and results, etc.) or intuitive mind maps based on the provided research topic. Once your outline is written and you have a good sense of your bibliography, GPT can assist researchers to manage it, including formatting in specific citation styles, or checking reference lists for accuracy. Consider the improvement that "find-and-replace" initially provided to editors. When we wrote by hand or with typewriters, we had to do this activity manually. Find-and-replace eliminated a lot of tedious work, but it took some preparation to make sure that

it replaced exactly what you wanted, and nothing more. With GPT you can now be very specific by asking in the way that you would ask another human: e.g., “Please find for me all instances in the following text where an Oxford comma is missing and add one where appropriate. However, do not do this for quoted text, which I would like to remain verbatim.”

That’s all great but there are some drawbacks to this efficiency. While you may indeed be more productive, you will lose some control over the creative and academic process. Take the recent example of how research has changed since the spread of digital PDF documents. With PDFs we may have become more efficient, but we are less open to serendipity. Will the next generation of scientists forever lose that beautiful feeling of sitting in a quiet library and thumbing through paper journals? If we get AI implementation wrong, we are at risk of ceding control to AI over what we write and publish.

There is also the real fear that AI will not only improve our writing, but also our laziness, and we’ll lose the love of discovery and life-long learning. That would certainly be a pity. Moreover, I believe there is the real risk that—between the creation of scientific content, its dissemination, and the synthesis of that content by other people—we end up in a situation where there is AI between most communications. Will our true intentions be meaningfully communicated when filtered through AI? It would be a bit like how I can “speak” Mandarin to someone when filtered through Google Translate. Or will there be instances where different AIs communicate with each other directly, and perhaps even disagree. A tiny example of just such a thing occurred to me while redacting this editorial. The GPT-generated output used in the following section did not insert a serial (Oxford) comma in a list. Microsoft Word’s grammar AI flagged this as a possible error. I tend to agree with Word in this case, so I took the final decision and added the comma. For the time being, the default relationship between humans and artificial intelligence should be like so: AI generation with human verification\*.

### The future of artificial intelligence

Let’s conclude by taking a broader perspective. As pointed out by historian Yuval Noah Harari, AIs are still in their infancy but have the potential to evolve rapidly, and unexpectedly, surpassing human capabilities in just a few years or decades, eventually reaching Artificial General Intelligence (AGI). AI is the first technology that can *make decisions* by itself, taking power and responsibility away from humans. Crucially, it has the potential to destroy or diminish democratic conversation because, while dictatorships may rely on one person or entity, democracy succeeds or fails via debate over information.

The challenge then remains to make the best decisions to mitigate risk while taking advantage of AI’s many benefits. Despite the risks associated with it, AI is already so potent that we cannot consider turning it off. To highlight my point, with this editorial in mind, I asked ChatGPT the following:

*Can you give me a bullet point list of 10 ways in which generative AI will be a benefit to the medical sciences?*

To which it replies:

*Certainly! Generative AI has the potential to significantly benefit the field of medical sciences in various ways. Here are 10 key benefits:*

**1. Drug Discovery and Development:**

*Accelerating drug discovery by predicting potential drug candidates.*

*Designing molecules with specific therapeutic properties.*

**2. Personalized Medicine:**

*Tailoring treatment plans based on a patient’s unique genetic and medical profile.*

*Generating patient-specific treatment recommendations.*

**3. Medical Imaging:**

*Enhancing image resolution and quality for more accurate diagnoses.*

*Automating the analysis of medical images for faster and more precise results.*

**4. Data Augmentation:**

*Generating synthetic medical data to augment limited datasets for research and training.*

*Creating diverse data samples for training machine learning models.*

**5. Disease Prediction and Early Detection:**

*Developing predictive models for early disease detection based on patient data.*

*Identifying patterns and biomarkers for disease risk assessment.*

**6. Natural Language Processing (NLP):**

*Automating medical documentation and transcription tasks.*

*Extracting valuable insights from unstructured medical texts, such as clinical notes and research papers.*

**7. Drug Repurposing:**

*Identifying existing drugs with potential applications for new diseases.*

*Reducing the time and cost of bringing drugs to market.*

**8. Genomic Sequencing:**

*Generating synthetic genomic sequences to study genetic variations and mutations.*

*Predicting the impact of genetic variants on health.*

**9. Treatment Optimization:**

*Recommending personalized treatment plans by considering patient history, genetics, and current health status.*

*Adjusting treatment strategies in real-time based on patient responses.*

**10. Medical Education and Training:**

*Simulating medical scenarios for training healthcare professionals.*

*Creating interactive educational content and virtual patient simulations.*

*Generative AI holds promise in revolutionizing medical research, diagnosis, and treatment by harnessing its ability to generate data, predictions, and insights, ultimately improving patient outcomes, and advancing the field of medicine.*

Curiously, one astonishing thing about this new technology is

\* How AI was used to create this article: I used ChatGPT to help craft an outline for which to structure my ideas; I used it to generate text to demonstrate its capabilities (“list of ten key benefits of AI to the medical sciences”).

that it took about as much time for me to type the query as it did for GPT to generate and output the results. As medical researchers, you are better equipped than me to judge the veracity of that response.

In this article I have mainly focused on point number six, Natural Language Processing. However, the overall benefits of AI are much greater than NLP alone. By now, you have most likely heard about Google's AlphaFold, which now can "predict the shape of a protein, at scale and in minutes, down to atomic accuracy." It has already classified over 200 million protein structures, including "nearly all catalogued proteins known to science." This freely available dataset <sup>[4]</sup> combined with CRISPR gene editing will allow enormous and unprecedented advances in the medical sciences. Separately, Elon Musk's company Neuralink has recently received regulatory approval to begin human trials <sup>[5]</sup>. Neuralink will precisely implant tiny microchips and wires into the brain with a surgical robot. This brain-computer interface allows acquisition and amplification of brain signals. Ostensibly it will help the paralyzed walk, but what else will it be used for? The AI dilemma <sup>[6]</sup>, then, is that we cannot now abandon AI technology. How could we, as they might help find a universal cure for cancer or other diseases, not to mention provide climate solutions, or offer an end to human labor. But can we control it? Much weight has been given to the possibility (often portrayed in science fiction) that AI will seek to dominate us. But this is a misplaced fear. Fear is an emotion, and AI doesn't (yet) have emotions. Rather, the present danger is that the goals of AI will not necessarily align with our own—the so-called alignment or control problem. To wit, consider AI researcher Eliezer Yudkowsky and philosopher Nick Bostrom's paper clip maximizer analogy, which hypothesizes a scenario where a future AI, programmed to stop at nothing to produce paper clips, would literally attempt to use every atom in the Universe (including those in human bodies) to do so. While this is a fantastical thought experiment, it illustrates the concern that AI algorithms will need to be somehow contained or programmed to act for the benefit of humanity, while not being given the capability to modify that instruction—a very non-trivial problem.

To be clear, in the short term, the power of algorithms to shape our beliefs and opinions is much more worrying than the fantasy of an AI takeover. Despite our intelligence, society is facing two major threats—ecological collapse and technological disruption—and instead of uniting to face them, we are dividing and fighting each other, with an observant AI along for the ride. To cite Harari again, "one of the biggest challenges facing humanity now is our inability to cooperate; if we can cooperate, we can deal with AI—if we can't it's hopeless."

There is no doubt that AIs will replace many jobs. But they will also create many new opportunities and capabilities. A recent study <sup>[7]</sup> carried out by researchers at the University of Maryland School of Medicine has reportedly found that ChatGPT provided correct medical information for the majority of the times it was queried. In the study, published in *Radiology*, three radiologists evaluated its responses and found that 22 of the 25 responses were appropriate. Two of the rejected responses were discarded

due to inconsistencies because of ChatGPT's tendency to give a newly generated answer each time. One response was discarded due to outdated information. This should be expected; not because GPT has any actual understanding of a subject, but because it was pre-trained on an immensely large online dataset that presumably would have mostly accurate information—for highly opinionated or subjective topics, such as politics or musical preference, one might expect that it does not give such accurate or even straightforward answers.

When it comes to education, ChatGPT-like tools have the distinct potential to replace most scholastic material. If we get it right, that style of personalized teaching would quite possibly enrich education. As with the introduction of the calculator into mathematics classes, we need to accept that students will be using GPT (and other Gen AI tools), and we need to design our evaluations accordingly. For instance, one might consider giving students the task of writing an essay *using* ChatGPT, then, back in the classroom, without AI, they could be taught to critique, improve, or modify it in a way that is deeply didactic.

We are still operating our education systems with an industrial revolution mindset, preparing students for "the workforce." But things are starting to change. Already, at my son's high school each pupil has a personalized Chromebook where they are encouraged to do their homework. As additional homework we use the wonderful app Brilliant <sup>[8]</sup> together. It's visual, intuitive, interactive, and non-judgmental. In addition, OpenAI, the creators of ChatGPT, have just announced a partnership with Khan Academy <sup>[9]</sup> to pilot a virtual GPT-4 tutor.

For the foreseeable future, AIs won't be able to fully replace the care and compassion that is so natural for humans, especially where it is so important, as in medicine. Similarly, while AIs may be able to communicate valuable information and knowledge (i.e., to teach us), they will not substitute the deep passion and inspiration of an effective teacher. However, we must only assume two things for even that status to be disrupted: 1) that intelligence is in fact substrate independent (i.e., it can manifest both in organic brains and in silico) and 2) that optimization improvements in artificial intelligence will continue at some rate to eventually arrive at super-intelligence. If we successfully navigate the challenges that lie ahead, we can reap the rewards and make the world a better place.

## References

1. <https://www.samharris.org/>
2. <https://ninaschick.org/>
3. <https://autogpt.net/>
4. <https://alphafold.ebi.ac.uk/>
5. <https://www.bbc.com/news/technology-66865895>
6. <https://www.youtube.com/watch?v=xoVJKj8lcNQ>
7. <https://www.medschool.umaryland.edu/news/2023/UM-School-of-Medicine-Study-Finds-ChatGPT-Helpful-for-Breast-Cancer-Screening-Advice-With-Certain-Caveats-.html>
8. <http://www.brilliant.org/>
9. <https://openai.com/customer-stories/khan-academy>