Large Language Models: the new actors for knowledge access

Summary:

Since the launch of ChatGPT in November last year, the popularisation of artificial intelligence is a fact. Its availability and applicability to the everyday life of the general public is a major difference compared to the arrival of other technologies that also marked a turning point in the evolution of mankind.

Behind this new technological milestone are technical issues in a wide range of fields covering economics, society, politics, security, education, industry and even climate. Having the technology at one's disposal is a bulwark of scientific capabilities, giving rise to the race that is taking place, which is discussed in this paper.

Key words:

Artificial intelligence, technological competence, knowledge, information

Large Language Models: los nuevos actores de acceso al conocimiento

Resumen:

Desde el lanzamiento de ChatGPT, en noviembre del año pasado, la popularización de la inteligencia artificial es un hecho. Su disponibilidad y aplicabilidad al día a día del público general supone una importante diferencia con respecto a la llegada de otras tecnologías que también supusieron un punto de inflexión en la evolución de la humanidad.

Detrás de este nuevo hito tecnológico hay cuestiones técnicas de ámbitos muy diversos que abarcan la economía, la sociedad, la política, la seguridad, la educación, la industria e incluso el clima. Disponer de la tecnología constituye un baluarte de las capacidades científicas, lo que da lugar a la carrera que se está produciendo, que se analiza en este documento.

Palabras clave:

Inteligencia artificial, competencia tecnológica, conocimiento, información.

Introduction

It is just over a year since the arrival of ChatGPT as a general-purpose artificial intelligence tool. This is one of the products of what are known as large language models, which are artificial intelligence tools that, based on statistical processing of large volumes of information in the form of texts on a multitude of subjects, are capable of answering questions posed by users by establishing a question and answer dialogue.

The technology, which has been in development for many years, has become a service available to the entire population. It provides functionalities that can be used to cover needs that, until now, were either not possible to obtain or were obtained in other ways. Among many other options, it is possible to obtain answers to specific doubts by using the information collected and processed during the training of the system, while it is also possible for the user to provide specific information to carry out operations such as summarising or extracting the main ideas from texts, translating content into other languages or modifications such as transforming the style in which a text is written to adjust it to the way of writing of an institution or a person.

The general availability of the tool has led to a growing identification of applications, the result of the approach of users with more diverse profiles exploring the applicability of the new technology to the resolution of their specific problems. The specialisation of these profiles and their experience, from the accidentally curious to the most reputable academic, is also making it possible to detect the problems and risks that the use of these technologies may entail.

The enormous potential identified in this technology also makes it a vector of commercial interest, while the availability of the technology, its regulation and the consequences of its use are issues that states must work on, as there are many legal, regulatory, strategic and geopolitical factors associated with a globally available service that is currently being provided by private companies.

Generative AI reaches the general public

The speed with which ChatGPT reached 100 million users in less than two months set a new record for the number of users of a new internet service. The interest in the technology, since it became available to the general public in November 2022, exceeded

the capacity of the systems deployed to provide the service and many users had to wait before they could get an answer to one of their questions.

Seven months later, the first decrease in traffic on the servers was around 10%, which can be attributed to the end of the teaching periods of the vast majority of educational activities[1] , to a loss of interest on the part of users or to the migration to the paid versions and the use of alternative access routes, such as Microsoft's Bing search engine or directly through the automated APIs that the company OpenAI makes available to developers[2] .
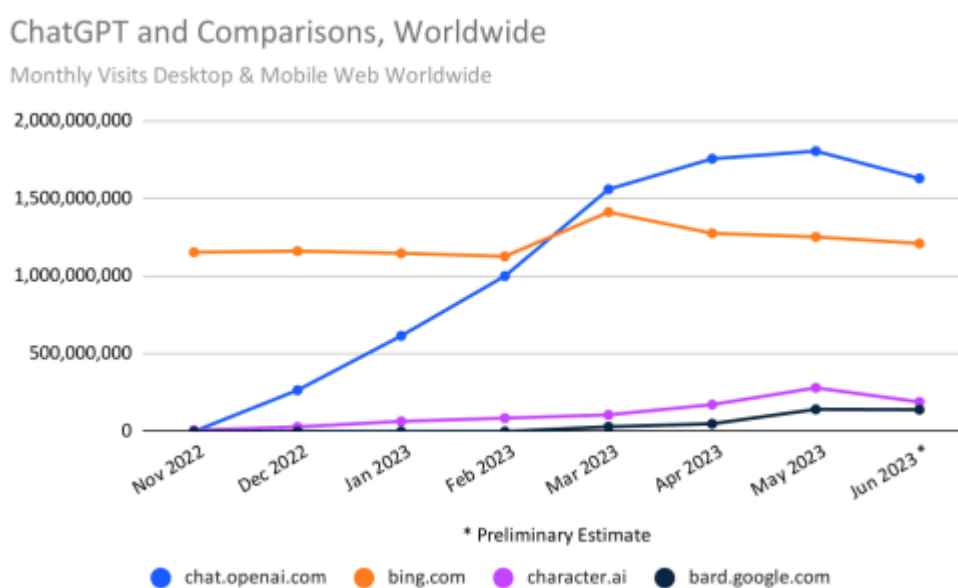

Figure 1 Evolution of visits to LLM services from mobiles and computers (Source: similarweb.com)

ChatGPT is the spearhead of a phenomenon, that of Large Language Models, which has given a real boost to the popularisation of artificial intelligence. Its success has forced the rapid deployment of alternatives in a race for positioning in a potential market with such a high latent demand. Google launched Bard as an alternative to ChatGPT, based on LaMDa, and Meta has LLaMa, both of which are commercial alternatives to be expected among internet giants that must earn their market through competition.

---

[1] QUACH, Katyanna. Now that you've all tried it ... ChatGPT web traffic falls 10%. Available at
https://www.theregister.com/2023/07/07/traffic_to_chatgpt/
[2] CARR, David F. ChatGPT Drops About 10% in Traffic as the Novelty Wears Off. Available at
https://www.similarweb.com/blog/insights/ai-news/chatgpt-traffic-drops/

Meanwhile, in China, several companies advertised the launch of their own products[3] , Tongyi Qianweb of Alibaba Cloud, Tiangong of Kunlun Tech or Ernie of Baidu, in what represents, in addition to a commercial response for those areas where the origin of the technology is not a priority, a strategic response within the technological confrontation that is taking place between the United States and China. No aspiring technological hegemon can afford to lack a tool that another aspiring hegemon possesses.

Launching a new product with features as complex as an LLM is not affordable for all companies, let alone immediately. Microsoft has partnered exclusively with OpenAI[4] , developer of the GPT model, and is providing the functionality of the GPT-4 version through its Bing search engine, as opposed to the less advanced GPT-3.5 version, which is the one that led to the ChatGPT revolution.

Another way to take advantage of the revolution is to incorporate the tool's functionalities under the prisms of tailoring the system's capabilities to specific sectors or providing mechanisms to cope with the problems and drawbacks that may arise from the use of this technology.

The need for control and the complexity of its implementation are not a trivial issue when the training of the system is automated and its operation must respond to ethical, political or legal constraints that are difficult to interpret, the transgression of which can have serious consequences for the service provider and, more importantly, for the safety of the users.

In this sense, accelerated deployments can pose a serious risk to owners in the face of unwanted behaviour, as happened to Microsoft with the artificial intelligence-based conversational bot Tay, which had to be deactivated a day after its launch due to the racist and xenophobic results it began to generate as a result of its uncontrolled interaction with (clearly malicious) users in 2016[5] .

---

[3] CHENG, Evelyn. China's A.I. chatbots haven't yet reached the public like ChatGPT did. Available at https://www.cnbc.com/2023/04/28/how-chinas-chatgpt-ai-alternatives-are-doing.html
[4] OpenAI forms exclusive partnership with Microsoft to build new supercomputing technologies on Azure AI. Available at https://news.microsoft.com/es-xl/openai-forma-una-exclusiva-asociacion-con-microsoft-para-construir-nuevas-tecnologias-de-super-computo-en-azure-ai/
[5] Tay, Microsoft's racist and xenophobic robot. Available at https://www.bbc.com/mundo/noticias/2016/03/160325_tecnologia_microsoft_tay_bot_adolescente_inteligencia_artificial_racista_xenofoba_lb

Access to information

How information is accessed has been changing from the origin of writing to the present day. In recent years, it is possible to identify four routes that have been positioning themselves as the options of reference.

Social networks are currently the main access point to information in developed countries. They provide immediate information about what is happening and are gradually displacing the traditional media, such as press, radio and television.

However, they still retain an important relevance because they provide something that is more difficult to achieve in social networks: reliability. They also allow for something that is much more complicated in social networks, and that is the question of archiving information through newspaper archives, which, with digital content, are more difficult to manage and make available to society.
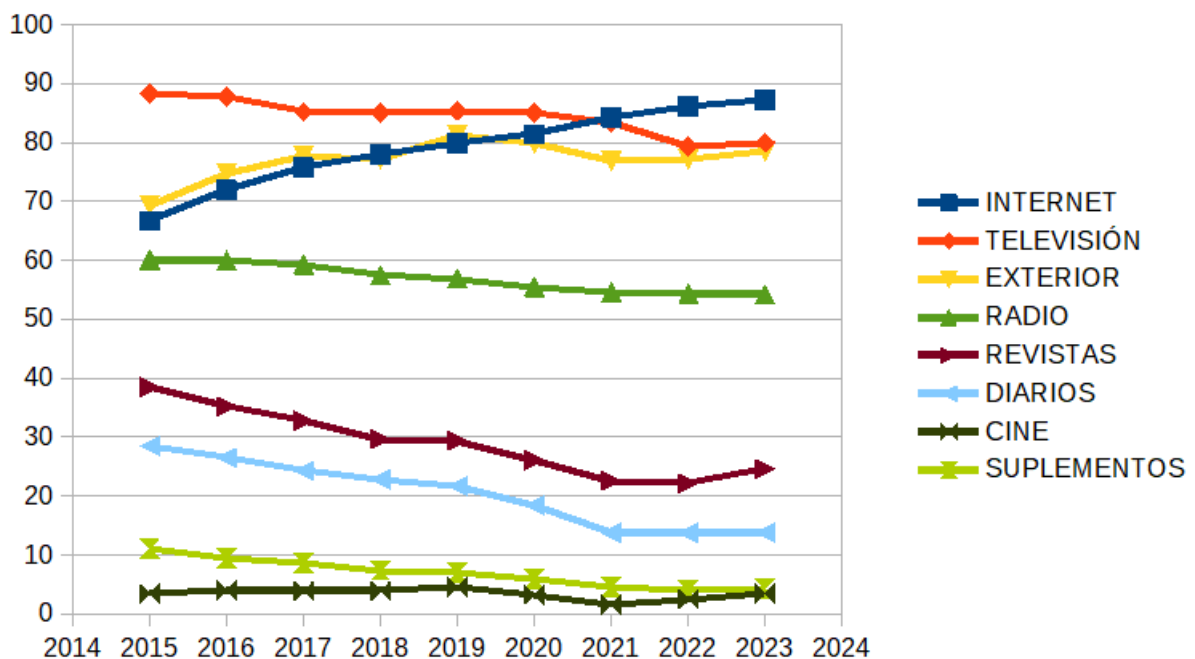

Figure 2 Evolution of media penetration (EGM AIMC data www.aimc.es)

There is no doubt that the internet still plays a predominant role when it comes to accessing information. This is the difference between the internet as it was used in the past and the new model of relationships through LLMs. Although in most cases it is not the source of the information from which work is done, it does provide references to specific content, publications and information. If there is one thing that is particularly remarkable about this model, it is infoxication, as the excess of information obtained by searching the internet on any subject is called. In practice, this is a problem because, at

present, the information available on the internet is openly accessible and there is a great deal of it. Its quality is, a priori, unknown, although the origin of this information (the website on which it is located) as well as the type of information in question (think of the content of a blog versus the website of an institution) allows the user to assess, to a certain extent, the reliability of the information.

In academia, traditional documentary sources such as books and periodicals, physical or online, are the reference sources. Both the publisher and independent bodies are responsible for assessing the quality of the publication through strict verification procedures, which is reaffirmed by meta-information such as the number of citations and the affiliations of the authors publishing in these sources.

This ecosystem has been evolving to incorporate new media, and the arrival of LLMs may mark a turning point in its evolution in a way that has never happened before. There is a growing trend among users towards free sources, especially if access to information through them is easier. In this, social networks stand out especially because the user becomes an almost passive consumer of the content that the algorithms select as most relevant to that user. The general information consumption model is thus covered.

For more specific queries, there is a migration from the use of content search engines to the use of LLM. Here there is a major shift because the user is no longer provided with data on where information of interest may be available, but the algorithm selects information deemed relevant and elaborates it in such a way that it is presented to the user as complete content that does not require further processing. The elaboration process filters the contents in an operation that eliminates information of dubious reliability for statistical reasons, but also eliminates the metadata that would allow the reliability of the information obtained to be assessed in the increasingly rare case that the user decides to check the information obtained.

Basic issues

LLM systems have been developed over many years by perfecting the "word embedding" technique, which consists of determining the word or words that should appear next on the basis of statistics and other algorithmic procedures. The advent of trained transformers, from which the initials P and T of GPT derive, marked a turning point that has been applied by many of the solutions currently available.

The operation of LLMs is the result of applying machine learning algorithms to knowledge corpora with which artificial intelligence is trained, from which a configuration of the system's operating parameters is obtained so that, by applying a request for information, a question or a statement to the input, processing its content, an output is provided in the terms that have been requested.

As indicated by the leader of an experiment in which the LLM GPT3 was subjected to pass the medical licensing test, it is a machine that does not generate knowledge and merely provides an output related to the content used to train the system. This mode of operation means that the system will not be able to consciously generate new information, but that the output will present a sequence of words that, ideally, corresponds to the response to the input request, limited by the precision of the information with which the machine has been trained and of which there is no way of certifying the precision, completeness and accuracy of the response, while it is also impossible to establish the relationship of this with the training information or the parameter relationships that have given rise to the response obtained.

This leads to various problems in the functioning of the system.

The first problem is the limitations of the system itself in terms of the maximum complexity it can handle, which will affect parameters such as the completeness and accuracy of the system. We are talking about billions of parameters that need to be adjusted through the training procedure.

Another problem lies in the system's interpretation of the text of the information request. Two questions with small differences between them may lead to results with different degrees of accuracy or completeness.

Today, given the extreme complexity of the system, it is impossible to establish clearly and precisely the criteria, data and information used to produce the answers. It is a black box that provides results under criteria of an eminently statistical nature by concatenating one word after another until whole paragraphs and texts are composed.

The system's training corpus is a vital factor since, as a result of the system's operating principles, any information or relationships that cannot be extracted from the information provided do not exist. The system will not be able to elaborate theories or conclusions that deviate from the knowledge contained in the training corpus. In fact, in some cases, the development of so-called special capabilities has been detected, as they seem to

generate results whose foundations or relationships were not evident in the input information[6] . In these cases, the system has been able to detect these less obvious relationships, which has been interpreted by users as the development of intuition or general intelligence on the part of the system.

On the other hand, there are the results considered as hallucinations, whereby the answers provided by the system do not correspond to reality. The system will generate an answer to the question asked and will select the content statistically from the information it contains. If it is asked, for example, which countries are members of NATO, it could generate a list with an arbitrary number of countries and these could be allies or not. This is correctly written text but the content may not be correct[7]. There is no mechanism to determine the correctness of the information and it is up to the user to establish the level of trust he or she places in it. This behaviour has come to be referred to as "stochastic parroting"[8] as it is a system that simply reproduces information without any judgement or knowledge about it.

In addition to training for system implementation, some of the solutions currently available can also learn from the content of users' previous requests for information. The dilemma arises as to whether this information can be used to solve other users' queries or only that of the user in question. Additional information provided by the user would contribute to the training of the system but could also lead to information leaks if sensitive and private information were to appear directly in the answers to other users.

The operating principles of LLMs are being criticised even by reputable artificial intelligence experts, as they consider that it is not possible to achieve AGI, Artificial General Intelligence, general artificial intelligence with the ability to think and reason like a human from an LLM. Its architecture does not lend itself to the understanding and elaboration of complex concepts but generates information in a statistical way[9].

---

[6] CLABURN, Thomas. Large language models' surprise emergent behaviour written off as 'a mirage' Available at https://www.theregister.com/2023/05/16/large_language_models_behavior/
[7] SMITH, Craig S. Hallucinations could blunt ChatGPT's Success. Available at https://spectrum.ieee.org/ai-hallucination
[8] BENDER, Emily M. and GEBRU, Timnit. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Disponible en https://dl.acm.org/doi/pdf/10.1145/3442188.3445922
[9] CLARK, Lindsay. Artificial General Intelligence remains a distant dream despite LLM boom. Available at https://www.theregister.com/2023/07/04/agi_llm_distant_dream/

Access and technology sovereignty

The barriers to access to certain technologies are high, the result of product development work in which the private component has been very important. This has led to a lack of visibility of the technology under development, which has meant that technologies that will become increasingly essential for state and private actors are currently out of reach even for states and international organisations, when sovereignty and strategic autonomy are prioritised by choosing local suppliers.

These barriers, given the existing ease and interest, which does not exempt the supplier from financial compensation, in implementing currently available solutions, are further enhanced not in terms of height, as the cost of developing alternatives in economic and time dimensions may be close, but in width or resistance, because the availability of solutions makes it difficult to motivate investment in mature stages of technology well behind the state of the art, and problems may also arise in relation to intellectual property, given the risk involved in developing equivalent capabilities when a solution is already implemented in the market.

In this last respect, it should be noted that the current solutions respond in most cases to a private business model backed by a private company with economic interests that are difficult to reconcile with the needs of states whose individual recruitment needs, if not to a drop in the ocean, it can be compared to the contribution of a river to the liquid mass of the planet. The power concentrated in these service providers is very great and the consequences have already been felt with similar scenarios when, following the implementation of data protection regulations in the European Union, one of the major social networks threatened to discontinue service provision in European territory. Recently, the sanction imposed on the same company for trying to circumvent data protection regulations on the basis of a self-interested interpretation of exceptions that were used to support the continuation of its activity was made public.

The perceived relevance of LLM means that states and supra-governmental organisations need to address the management of development, access and dependence on these technologies. As has been the case in recent decades for most technology-related issues, there has been a phenomenon of pushing technologies to be applied to existing problems.

There are even states that have either restricted their use or vetoed the use of technologies that are open to the general public. The reasons why several states have restricted or banned the use of LLM available online are diverse. These include countries such as Russia, China, North Korea, Cuba, Iran, Syria and Italy[10] .

Of the group, the presence of Italy, which vetoed access to the platform from its territory due to data protection issues, is very relevant. This veto put a large number of European users on edge because of the possibility that the arguments put forward in Italy could be used in other EU countries to veto access.

Ukraine was also on the list due to the impossibility of determining the use by citizens of the Russian-controlled territory of Crimea.

By way of example, OpenAI, the company that created ChatGPT, maintains a list of countries that can use the API to access its services and, for reasons of blocking, prohibition or omission, the following countries have been excluded:

- Afghanistan
- Bhutan
- Central African Republic
- Chad
- Eritrea
- Swaziland
- Iran
- Libya
- South Sudan
- Sudan
- Syria
- Yemen

---

[10] www.digitaltrends.com/computing/thiese-countries-chatgpt-banned/

Information leaks

In the information society, data are the main assets of individuals and organisations. Confidentiality thus becomes a basic security issue as endorsed by the European Union's data protection legislation.

Quite a few companies have limited or vetoed the use of generative tools because of concerns about preserving the confidentiality of information[11] . The list includes large technology companies, which a priori should be less distrustful of the digital world, as well as other companies where caution is reasonable because of national security concerns, such as Northrop Grumman, which works on defence projects.

When the system is requested to summarise or transform content, either by rewriting it in a different style or translating it into another language, it requires providing the system with the information to be transformed.

Systems can use information provided by users to fine-tune their operation. In this case, information provided by one user may be used indirectly to generate answers to another user's questions. Knowledge leakage may occur whereby the LLM uses information from one user to provide the answer to another user.

Incorrect system design can also lead to information leakage through a cyber-attack exploiting insufficient security measures that allow access to individual, contractually confidential data by third parties. For a few days, ChatGPT users' information was exposed due to a vulnerability in one of the libraries used for its deployment[12] . The vulnerability allowed users to access the chat information of other arbitrary users. All information entered in the chat with which one interacts with the system could appear in another user's chat. If text had been entered for summarisation or translation, the text could appear as it was in another user's chat.

The work of the prompt engineer, i.e. the expert in charge of generating the question in the right terms to get exactly the answer you want, can also be a source of information leakage. The expert's knowledge of the area in question is being used to construct the question. Although the operation of the system does not transfer this information to the

---

[11] MOK, Aaron. Amazon, Apple, and 12 other major companies that have restricted employees from using ChatGPT. www.businessinsider.com/chatgpt-companies-issued-bans-restrictions-openai-ai-amazon-apple-2023-7

[12] KOVACS, Eduard. ChatGPT Data Breach Confirmed as Security Firm Warns of Vulnerable Component Exploitation. www.securityweek.com/chatgpt-data-breach-confirmed-as-security-firm-warns-of-vulnerable-component-exploitation/

information processing of other users, the service provider receives the question adjusted in precise terms and can extract relevant information in relation to the required topic. Imagine a request for information on research in which one or more current lines of work in that field are specifically excluded. And to this can be added the metadata of the connection, which could be used for geo-strategic intelligence purposes.

In order to obtain higher quality answers, the user is being asked to ask contextualised questions. The introduction of context is a transfer of knowledge. LLMs are being educated to answer their own questions. But this experience, this knowledge, is being transferred to AI. You may not use it immediately for content generation for other users, because you may incur risks such as disclosure of sensitive information present in information requests (personal or confidential data) or bias due to the ideological leanings of the user who is accessing the LLM's capabilities. In the future, this information can be used for knowledge generation. It is a learning process of the system in which all users who make use of it become improvised teachers who are transferring their knowledge through the elaboration of questions that include those relevant data that are used for the particularisation of the answers.

The Kessler effect of cognitive mastery

Internet search engines crawl internet content in order to index it and propose it as an answer to a user's search for a certain keyword. Although the information acquisition, processing and prioritisation algorithms used to provide answers to users' searches are proprietary (known only to a limited group of individuals), it is possible to deduce questions about them. The ability to process openly available information, a necessary condition for internet content-sweeping systems to be able to access it anonymously, is limited. It is also evident that it takes some time to incorporate the newly available information into the information corpus from which the responses and the criteria for selecting and prioritising them are developed.

The web is the platform that hosts the general information from which general purpose LLM systems are working. The contents generated by individuals, organisations and collectives constitute the knowledge that the systems will analyse to adjust the parameters that will lead to the required responses.

Much of the information currently available on the Internet is also the result of the so-called Web 2.0 in which the user became a *prosumer*, a term which combined the role of consumer of available content with the role of producer of new content which was made available to other users through blogs and other means of collaboration.

The advent of LLM can give rise to a new source of information available on the internet if users start distributing the content generated. The training of new LLMs will take this information into account and a feedback effect will occur whereby new content will be generated from information that has already been elaborated from previous content. In the event that errors exist at some point in the chain, these will be perpetuated as more and more cycles of information processing convert them into statistically more relevant information. Avoiding this would require identifying automatically generated content to avoid its use in the training of new systems. However, the systems being developed to detect when content has been generated automatically or by a person have very unpromising accuracy and reliability rates for false positives and false negatives[13] .

This would produce an effect similar to the Kessler effect[14] in the spatial domain. This effect describes a chain reaction scenario whereby a space debris, upon collision with another object in orbit, generates more debris that will collide with more objects until the exponential evolution of the number of debris and collisions would make the use of the orbit unfeasible because it would be impossible to preserve the integrity of a platform located there.

The growing presence of information of dubious quality in the cyber domain, which can in turn be a source of training for new systems, would render them useless because of the impossibility of using them reliably and productively. Not being able to differentiate between true content and manipulated or outright false content, whether generated by artificial intelligence or not, would undermine the foundations of trust and the dissemination of information on the internet to the point of making it, in practice, useless

---

[13] ARNETT , Stephanie. This is how easy it is to fool AI-generated text detection tools. https://www.technologyreview.es/s/15532/asi-de-facil-es-enganar-las-herramientas-de-deteccion-de-textos-generados-por-ia
[14] KESSLER, Donald J., COUR-PALAIS, Burton G. Collision frequency of artificial satellites: The creation of a debris belt. Journal of Geophysical Research. Space Physics. Vol. 83 Issue A6, 1 June 1978. https://doi.org/10.1029/JA083iA06p02637

for any practical matter beyond the location of certain non-sensitive or easily verifiable information.

The effects of LLM on teaching

The ease of access to LLM tools by the general public has generated deep concern in the academic sector as it invalidates the mechanisms established for the training and education of students when used inappropriately.

Tasks such as writing papers on specific subjects, which involve a process of gathering information, reading, assimilating content and producing a document based on the information covered and the experience acquired by the learner, are no longer meaningful. All this can be done in a short interval of time by the LLM system. The learner only has to enter the terms under which the system elaborates the document and he/she will get the document according to these criteria. However, the desired product of this process is not the document as such, but the skills and abilities developed by the learner as well as the knowledge assimilated. The learner will have satisfied the obligations but will not have acquired the necessary skills.

The problem is not only limited to the lack of development of the learner's skills. The assessment of their performance, which requires objective tests to determine the acquisition of the required knowledge, also becomes a challenge. The purpose of assessment, especially in those academic fields in which a qualification is provided to enable students to carry out certain professional tasks, is to certify that the student has the minimum skills required to carry out an activity professionally. With the new educational models, which promote assessment by competences on a continuous basis, instead of traditional examinations or one-off assessment tests, tasks such as the one described are useless for this purpose.

In the field of education, especially in subjects that lend themselves to ideological bias, there are many risks involved. The operating principles of the systems can lead to the results obtained following trends, resulting in their amplification by the education system. Those ideas that are most often reflected on the internet will take precedence over positions that receive less support.

But this risk may be due to issues outside the limitations or characteristics of LLM and result from direct intervention on the information selection procedure. Mechanisms can

be put in place in the system so that the results it provides support or refute certain ideas. It can thus become a tool with which to transmit ideology through the education that students receive.

However, it is also essential to include these tools in educational cycles. There are tasks in which they are very useful, such as translating content or extracting the main ideas from long documents. Knowing how to use them correctly can increase students' performance. To do so, they must know how to access their functionalities, correctly identify those functions in which they can be of help or support and develop a critical sense that allows them to evaluate the quality and validity of the results obtained.

Environmental issues in the spotlight

Computer technology has been in the spotlight from an energy point of view since a computer initially had an electricity consumption comparable to that of a small city. Providing it with the power to perform its function was an engineering exercise comparable to the development of the system itself. Over the years, technological complexity has given way to concern about electricity consumption in both its dimensions, in terms of the cost of the energy required on the one hand, and also in terms of the climate footprint of energy consumption in terms of tonnes of greenhouse gases or equivalent.

The water consumption required for cooling data centres is also[15] under the scrutiny of the bodies most concerned with environmental impact. Blockchain-based algorithms, especially those based on proof-of-work, on which most cryptocurrencies are based, have drawn attention to the high energy and water consumption associated with high computational needs. Water is another factor that is also triggering red flags.

As a result of being a commercial product provided by a company, regardless of the fact that it can be used free of charge, the information available on the main activities of the system is scarce. In the absence of concrete and precise information, it is not possible to

---

[15] Hidalgo, Mar. Energy and water consumption in data centres: sustainability risks. IEEE Analysis Paper 69/2022. https://www.ieee.es/Galerias/fichero/docs_analisis/2022/DIEEEA69_2022_MARHID_Datos.pdf (accessed 15/05/2023)

develop metrics to determine the consumption associated with the operation of the system, both for training purposes and for the development of each of the responses.[16]

With the current sensitivity to the high energy consumption associated with new techniques based on massive algorithms, voices are also being raised to warn about the high consumption of language models. The use of the technology involves two distinct processes that share the characteristic of requiring very high energy consumption. Firstly, the training of the system, whereby the corpus of information is processed and the operating parameters are adjusted to achieve the final system, has been estimated to consume around 90 days for a population of 3,000 people.

Once the model is trained, each time a user makes use of the trained service, energy consumption is also required to process its input information, access the necessary data and elaborate the necessary response and transmit it to the user. This operation, which involves a consumption well below that required for training the system, is repeated for each user that makes use of the service, resulting in a growth in consumption directly related to the number of people, as well as the number of applications for which the newly available technology is used.

Solutions are beginning to emerge from experimentation in real situations where solutions of lower complexity could provide more appropriate solutions with lower resource consumption. On the one hand, by reducing the number of parameters used to control the operation of the network, which results in a generally shorter learning process, a simpler implementation and requires less energy to elaborate each response. On the other hand, also reducing the knowledge corpus from which the network is trained. A smaller corpus will provide a less particularised training (it will be more difficult to obtain different results for small variations of the input request), usually shorter in time and, with a correct selection of the corpus, can provide higher quality answers with a lower complexity of the system.

---

[16] Josh Saul and Dina Bass. Artificial Intelligence Is Booming-So Is Its Carbon Footprint. Available at https://www.bloomberg.com/news/articles/2023-03-09/how-much-energy-do-ai-and-chatgpt-use-no-one-knows-for-sure

Conclusions

The impact that the arrival of the major language models is having stands out both for the speed with which their functionalities have been incorporated into a multitude of tasks and for the potential attributed to them to modify the day-to-day life of citizens.

There are many areas where there is an awareness of the risks that technology can pose to privacy and information security. However, in the wider society, issues such as these are overshadowed when contrasted with the possibility of free use of technology that provides answers to needs that until recently did not lend themselves to another automated solution.

The business world is responding by limiting the use of these technologies among employees, given the risks of information leaks that can affect business. LLMs (broadly understood as general-purpose work tools) are thus kept out of companies' supply chains. Despite service delivery agreements and legislation protecting information, there is mistrust about the proper preservation of information that may be sensitive when it leaves company borders. However, companies in the sector are campaigning intensively on the benefits and functionalities that the use of these technologies can provide to increase productivity, reduce costs and develop products that were unthinkable until very recently.

Internationally, we can also observe the reserve with which the different actors are considering the tools developed by third states. The creation of multiple tools for similar purposes responds to various needs such as security of supply, confidence in the technologies, control over the tools and, also very importantly, the importance of providing this type of technologically advanced systems to third countries in economic, strategic and geopolitical terms.

As is the case with the implementation of what are known as new technologies, the complexity and specificity of the technology means that there is a general lack of knowledge, resulting in a lack of the necessary precautions to provide users with a secure service. The need assumed by all types of actors to incorporate technologies even at an untested stage, given the risk that late access will make it impossible to play a relevant role in a fiercely competitive environment, is resulting in people without sufficient knowledge of the capabilities and risks having to make decisions. Also the fashion effect and the image effects that the incorporation of this technology, which looks set to become the new internet, brings to the dialectic between companies and institutions and their

clients and users, are fuelling the exponential growth that has taken place in such a short period of time.

The transformations to which the widespread implementation of these technologies may give rise to are accompanied by new risks. All levels of society will be involved in the new solutions, and it is therefore necessary to establish mechanisms to control the new scenarios in order to avoid undesired or catastrophic situations due to the materialisation of risks that are continuously being identified in view of the opportunities that are opening up in the future.

David Ramírez Morán*
IEEE Senior Analyst
@darammor