LARGE LANGUAGE MODELS LIKE CHATGPT AND THE WORKPLACE

Artificial intelligence in the form of ChatGPT was a massive and almost instant phenomenon. One year ago – in November 2022 – OpenAI released its ChatGPT model on the internet. Within only a few days, the model had been tried by more than one million users. The release of this chatbot caught OpenAI's competitors off-guard, and many companies had to scramble to get competing products launched.

Unfortunately, the media hype generated by the rapid uptake might have caused some decision-makers to ignore paying more attention to the rapid developments in the generative artificial intelligence landscape.

It would be a mistake for decision-makers to allocate the developments in the AI field to the IT department. These developments can change how data is used and decisions are made. Even the current AI tools enable manufacturers to unlock their data and integrate diverse business areas like logistics, manufacturing, customer service, and more.

TECHNOLOGICAL NOVELTY OR MERIT

ChatGPT is an example of a Large Language Model (LLM). It is a form of generative artificial intelligence that uses a statistical language model trained on massive amounts of data.

Since this release, OpenAI and many other competitors have released more powerful models. While some LLMs are directly accessible by consumers, like ChatGPT, Google's Bard or Anthropic's Claude chatbots, the actual disruptions are happening behind the scenes at the level of cloud computing and processing large datasets. For instance, the latest iterations of OpenAI's application programme interface (API) allow developers to create their own AI applications at a minimal cost. These APIs have already started to change many desktop and smartphone applications like Onedrive, Dropbox, and Grammarly, as well as photo and video editing software. While we see the front-end, the real power is at the back-end of these applications.

However, we are still in the early days, with the ripple effects of these technologies expanding into more and more domains.

Many of the disruptions caused by Large Language Models have been due to the experimentation by thousands of open-source developers competing for investors trying to catch up with this new wave. The clamour for attention is further fuelled by larger companies desperate to license code modules that they cannot afford to develop from scratch.

We are still in the early days, with the ripple effects of these technologies expanding into more and more domains.

TECHNOLOGICAL CHANGE AND INNOVATION SYSTEM OBSERVATORY

The aim of the Technological Change and Innovation System Observatory project is to support the Department of Trade, Industry and Competition (the dtic) and industry sectors to develop an integrated, strategic response to discontinuous technological change and disruptive innovation. It aims to equip public and private organisations to become more sensitive to global technological shifts, and the changing demands placed on the innovation system, the manufacturing sector and its stakeholders.

Author:
Dr Shawn Cunningham

Contact:
sc@mesopartner.com
TIPS
info@tips.org.za
www.tips.org.za

November 2023

In brief, LLMs allow companies with large datasets to develop AI applications to leverage their data at a fraction of the cost of what it took just two years ago. However, LLMs are also changing how new visual or audio data is processed, generated or edited. It is, therefore, important for decision-makers to understand where their operations might be vulnerable to disruption by competitors that figure out how to harness these powerful algorithms to create new content or value at a significantly reduced cost.

One of the most powerful uses of LLMs is to identify patterns in large datasets that may be disconnected at present. This can be achieved by uploading diverse company data and then investigating patterns. This is not an information and communication technology (ICT) function but a business intelligence function supporting decision-making and strategy, enabled by data science and ICT skills.

NEW COMPETANCE REQUIREMENTS

A challenge is that while expertise is available to larger manufacturers and corporations, smaller firms might have to figure this out by themselves, as ICT expertise may be expensive on a pay-by-use basis.

The simplest way to access the features of Large Language Models like ChatGPT is through a browser. However, the functions of LLMs have already been absorbed into many desktop and smartphone applications. Therefore, the competence requirements to use basic features are going down.

However, to get real value from an LLM via a chatbot, a user must be able to formulate prompts or provide instructions. For instance, instructing a chatbot like Bing to generate an abstract of a document is not as simple as just uploading a document. The user must provide clear instructions for style, emphasis and format.

A more advanced form of using LLMs is to string together different LLM applications to create novel content. For instance, a user could ask ChatGPT3 to write a movie plot and to develop characters. These outputs can then be fed into Midjourney.com or OpenAI's Dall-E to create characters or animations while using another interface to generate a soundtrack or a voice-over. To make this work, the user must determine how to select different applications and transfer data between different applications to get the desired result.

An advanced use of LLMs is using dedicated AI applications on a desktop or a smartphone. For instance, many companies using Microsoft 365 subscriptions have access to potent AI software. Others, like Amazon Web Services, Adobe, IBM's Watson, UiPath and many more have powerful data analytics and automation applications. These applications allow companies to develop their applications drawing from libraries of functions, with the processing taking place using cloud computing. To use these, a user must understand data architectures, information management and managing cloud processing systems. Even though the functions are easy to use, coding languages like Python and basic HTML become necessary at some point, especially if front-end applications must be developed or if data must first be processed locally.

When a company wants to use its own data, process design, process integration, data collection, data security and protection of intellectual property, management competencies become more important. These capabilities are often beyond the competencies of IT personnel providing support and maintenance in an office environment. A challenge is that while these kinds of expertise are available to larger manufacturers and corporations, smaller firms might have to figure these out by themselves, as ICT expertise may be expensive on a pay-by-use basis.

SKILLS AND KNOWLEDGE

Many new applications launched by open-source developers are easy to use by self-trained or ICT literate users. The applications released by larger developers typically come with manuals, online courses or support communities.

However, for individuals who are not computer literate, or who are not used to figuring out how to use the features or functions of new software, using LLMs productively is not so easy – unless the user interface is easy to understand.

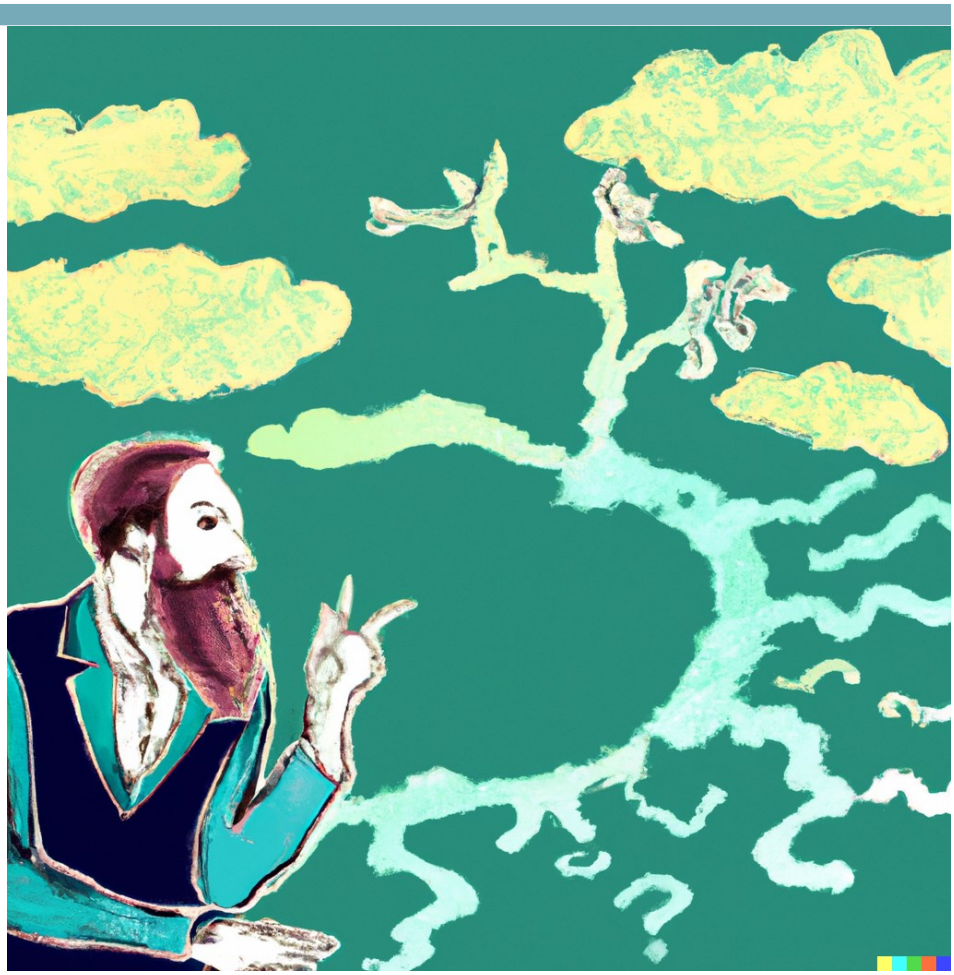
Developing AI solutions or harnessing the features of the software released by large vendors would typically require an understanding of either a functional domain (like industrial engineering, graphics design or life sciences) or experience in working with large datasets and cloud computing interfaces.

There are several different levels of skills.

	FUNCTIONS	EXAMPLE	SKILLS SOURCE
Digital literacy combined with functional expe-rience	Using LLMs through an application, browser or app	Using a chatbot to improve text or to generate an image	Self-taught, learning from others, social media
Intermediate digital literacy	Using specialised applications to generate high-quality outputs	Using Photoshop's AI functions to manipulate images, combining different LLMs for different functional purposes	Self-taught, learning from online tutorials or vendor instructions, online training programmes
Expert skills	Developing own applications, using own data in cloud processing applications Managing the different specialists that are needed to develop new solutions	Creating data architecture on Amazon Web Services or Azure, uploading data sets and designing process and automation flows	Digital expertise/ qualifications in conjunction with other qualifications in finance, data sciences engineering or other domains
Advanced skills	Data encryption, data integrity management, coding of customised applications	Using LLM architecture to develop in-house applications	Data sciences qualifications combined with software development expertise

For individuals who are not computer literate, or who are not used to figuring out how to use the features or functions of new software, using LLMs productively is not so easy – unless the user interface is easy to understand.

Currently, the LLM technologies are developing at an astounding pace. However, the foundational qualifications of software development, user-design interfaces, data science and information management qualifications are still an advantage.



This image was generated using Dall-E. The prompt used was "ideas circling the mind of a philosopher".

LLMs are very different to many preceding technological changes because individuals and employees can make decisions about adopting certain tools without management even being aware of these decisions.

ORGANISATIONAL ARRANGEMENTS

At an organisational level, there are again three levels of management.

- The first is to manage how users draw on LLMs, and how the risks are managed. For instance, when users upload proprietary data into an open-source application, a company might inadvertently share their intellectual property. Or when a company uses a chatbot to generate a media statement, undetected inaccurate information might lead to legal repercussions or reputational damage.
- The second level of management is to reimagine how AI tools like LLMs change the way different business functions can be re-connected. AI tools enable better decision-making, and can identify patterns or opportunities by drawing together different disconnected datasets.
- A third level is the decision by management to absorb some AI tools into their business model. To stay relevant and abreast of new developments is like stepping onto a fast moving escalator. Beyond the original commitment to integrate AI technologies into processes, companies would have to put in place mechanisms to remain abreast of relevant developments. When new developments become available, or when enhancements to existing technologies are possible, these will most likely not only affect ICT systems, but it may require more frequent assessments of business models, how an enterprise is organised, or even the strategy of the enterprise.

At the same time, LLMs are very different to many preceding technological changes because individuals and employees can make decisions about adopting certain tools without management even being aware of these decisions.

For instance, a business analyst could use ChatGPT to increase their productivity by a huge margin by simply using a web browser interface. If management does not encourage the use of these tools, companies might lose out on huge productivity increases at the level of individuals or small teams. As users increasingly use these new digital tools to enhance their work, it might free up valuable human potential that could be allocated to performing new functions or solving different problems. This is a hidden revolution that is unfolding in many work areas.

VALUE CHAIN EFFECTS

With the widespread dissemination of technologies like LLMs, we can expect disconnected functions in value chains to become more interdependent. This tighter integration might make reconfigurations in value chains possible because AI tools like generative AI solutions perform best with large and unstructured data sets.

MARKET NOVELTY AND NEW FUNCTIONS

LLMs can be expected to introduce changes in many different parts of value chains, connecting the data and decisions made by users with the whole production network. We can expect to see consolidation in many value chains while at the same time seeing new kinds of knowledge-intensive business functions emerge.

For example, a jewellery designer and manufacturer developed their own ChatGPT interface that allowed clients to design their own jewellery online. Once the client is satisfied with the design, the system automatically integrates the structural design elements, and then generates a price estimation and an expected delivery date. This interface required the manufacturer to integrate previously loosely connected functions, and it also required integrating the systems of suppliers, logistics partners and manufacturing process management into one system.

STRATEGIC IMPACT ON FIRMS

Many of the skills described earlier are about ICT and data management skills. Yet, the most important competence is the ability of managers to understand how AI integrates data and distributed processes in new configurations and how new digital technologies challenge or change the possibilities of how workplaces and organisational structures can be reimagined. These changes go beyond a single company to their supplier networks, markets and also their existing and potential competitors.

The fast pace of developments requires that senior management pay attention to the potentials and risks of these new technologies and that sufficient resources are allocated to trying new technologies in contained developmental sandboxes. Companies can gain experience in new technologies without disrupting existing processes by experimenting and learning about new technologies in sandboxes. Again, understanding the potential of these new technologies is not a typical ICT function, it requires an understanding of both the whole organisation and its function as well as the possibilities offered by a range of new technologies.

Furthermore, South Africa is far from the developments. Therefore local companies would have to closely monitor the emergence of substitute business models or the potential of new competitors moving from other sectors in their local or regional markets.

The fast pace of developments requires that senior management pay attention to the potentials and risks of these new technologies, and that sufficient resources are allocated to trying new technologies in contained developmental sandboxes.

REGULATORY REQUIREMENTS

Pressure is increasing on governments to regulate the development of certain AI technologies. The US government also regulates the trade of processing technologies with China. There are also global calls to protect intellectual property rights, as many developers of generative AI technologies have used published books, articles and online material to train their models.

Furthermore, LLMs trained on publicly available data have also been weak in overcoming bias and for expressing discriminatory remarks. Given the sensitivity about race, politics and gender in South Africa, companies using LLMs in their customer engagement or document management processes would have to take extra steps to ensure human oversight.

TECHNOLOGICAL CAPABILITY IN SOUTH AFRICA

Pressure is increasing on governments to regulate the development of certain AI technologies. There are also global calls to protect intellectual property rights, as many developers of generative AI technologies have used published books, articles and online material to train their models.

Although South Africa has good fibre infrastructure, and a number of high performance data centres, the country is lagging countries like Brazil in local AI processing capability.[1] However, several data centres in South Africa use the previous generation processors. These centres are certified to global standards and already host cloud computing firms like Microsoft, AWS, Google and others.

Historically, South Africa has been involved in cloud computing for more than a decade, with the early competence that became Amazon Web Services originating in the Western Cape. There is a concentration of digital media companies in Cape Town that are already using the latest AI-powered animation and media processing capabilities. In Gauteng, a concentration of finance, insurance and retail corporations are already using AI-technologies.

Cirrus is leading a consortium called the AI Initiative, bringing together academia and industry partners to campaign for support to establish a local processing hub, or to at least get access to the current processing technologies. The amounts required to establish this presence is in excess of US$300 million. The consortium is concerned that South Africa risks falling further behind the global AI-compute race that is unfolding, especially as high-speed processing enables scientific research and drug discovery. The challenge is not only the lack of the required processing infrastructure, but also the lack of skills in managing these kinds of systems. In this consortium, the University of the Witwatersrand is leading with exploring the potential of AI-technologies in scientific and medical research.

Many universities offer programmes in data sciences, information management and programming. However, only a few programmes are focused on AI and the management of large datasets. For instance, Prof Emma Ruttkamp-Bloem of Pretoria University is active globally on issues around the ethics of AI. The Department of Communications and Digital Technologies, the Tshwane University of Technology and the University of Johannesburg launched the National AI Institute in November 2022. This institute is focused on research and development, and its early results showed novel applications of AI in agri-processing and other areas.

Various programmes at the CSIR have digital technologies integrated into their portfolios. For instance, the NextGen cluster supports e-government, operational intelligence and networked systems. Another cluster is focused on smart cities. The CSIR's AI and Extended Reality research team uses machine learning for predictive modelling and sensory systems. The CSIR and the University of Pretoria are also participating in regional and international working groups on topics of bias, governance and the ethics of AI.

[1] *https://liftrinsights.com/cloud-regions-map/*