What are Large Language Models ?

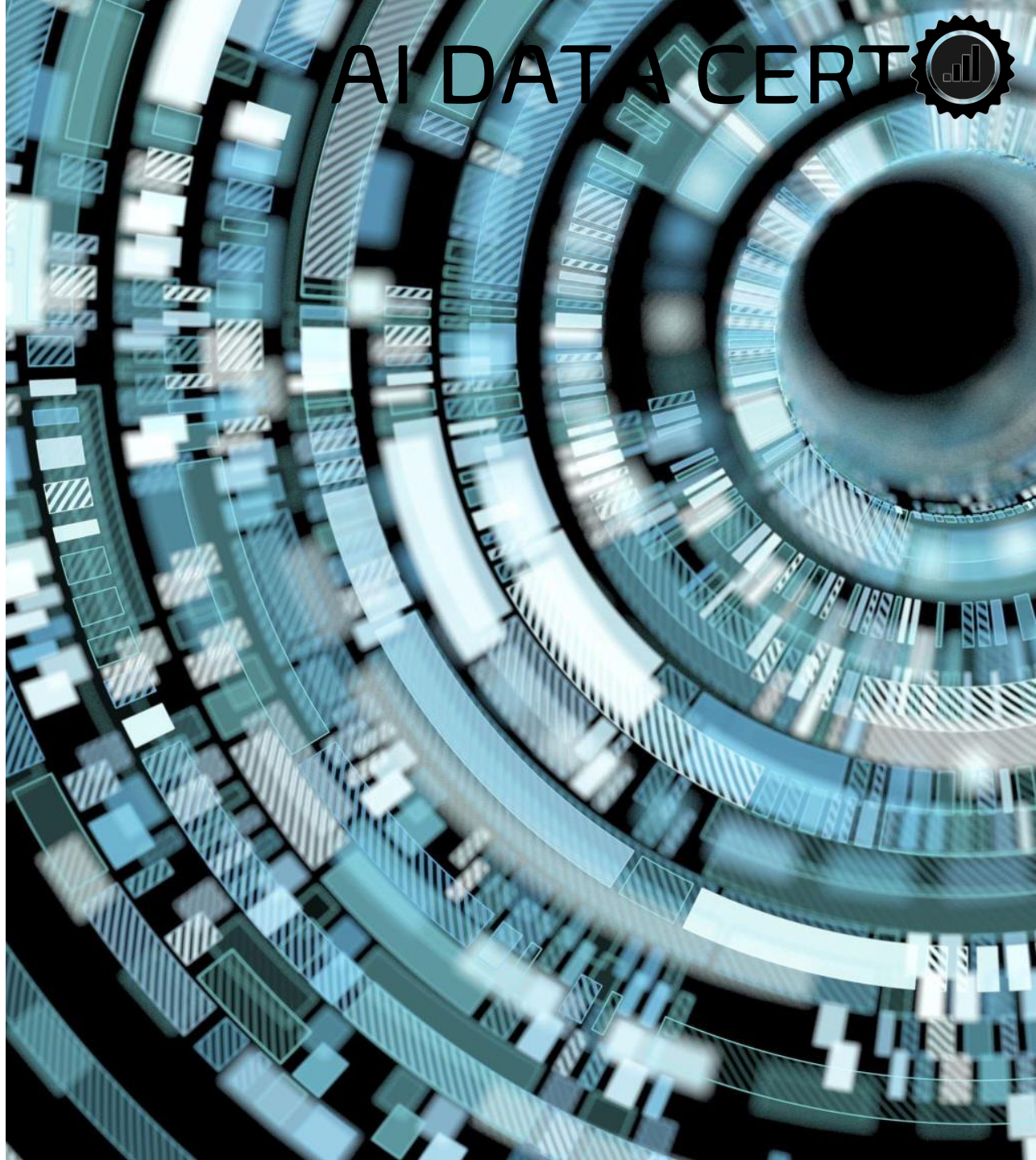
AI DATA CERT

September 2024

Introduction

Large Language Models (LLMs) are a type of artificial intelligence (AI) designed to understand and generate human-like text. They are used in various applications that many of us interact with daily, such as chatbots, virtual assistants like Siri or Alexa, and even in tools that help write emails or suggest text as you type. These models can process and produce text that sounds natural, making them increasingly valuable in both personal and professional contexts.

Whilst LLMs existed well before 2022 they became popular to the masses with the release of Chat GPT 3.5 in November 2022. This popular chat bot enabled users to use both its understanding and generation of human-like text to benefit from many efficiency gains in their daily lives, and since then model development has seen exponential increases in investment and in turn model performance.



How do these models work?

LLMs are built using three essential components: data, computing power, and a powerful algorithm. The combination of these three elements explains why LLMs have made such a significant impact, particularly since 2022.

Data: Over the past two decades, we have been generating and uploading vast amounts of data online. This includes everything from social media posts and articles to entire websites. All of this information is stored on servers worldwide and accessed via the internet. This massive volume of text data serves as the "knowledge base" from which LLMs learn.

Compute Power: Alongside the growth in data, there have been significant advances in computing technology. One key development is the use of Graphical Processing Units (GPUs). Originally designed for the gaming industry, GPUs have revolutionised how quickly and efficiently computers can train these models. The improvements in speed and cost-effectiveness have made it feasible to train LLMs on the vast datasets available today.

Algorithms: The third key ingredient is the algorithm, which is essentially the brain of the model. Recent advancements in computer science, particularly the development of the transformer architecture, have enabled LLMs to process and generate text more effectively than ever before. The transformer architecture allows these models to understand context and relationships between words in a way that was not possible with earlier methods.

Learning Process: LLMs learn to understand and generate human-like text by analysing as much online data as they can access, often combing through billions of web pages. The learning process involves predicting the next word in a sentence based on the words that come before it. The model makes a prediction, compares it to the actual word, and then adjusts its internal parameters to improve future predictions. By repeating this process billions of times, the model gradually learns the meaning of words and how they fit together in different contexts.

Example

Consider the sentence "Barry the dog was ..."

The model might predict that the next word could be "running," "barking," or "hungry." These are all plausible continuations based on the limited context. However, if we add more context:

"Barry the dog had just finished following his owner's scent for miles upon miles and was ..."

With this additional information, the model might predict words like "exhausted," "tired," or "hungry." The added context helps the model choose words that are more appropriate for the situation. This ability to use context to make better predictions is a key strength of LLMs, and it’s why they can generate text that feels natural and relevant.

At the end of this process of advanced next word prediction there is a process called Reinforcement Learning from Human Feedback. Here the next word prediction model is tweaked by real humans to make the responses more accurate, more human like and also safe.

Later on as you learn more about how to use LLMs you will hear advice to suggest that you add as much context to your prompt as possible to achieve a more specific answer. Hopefully you can see with the example above why this works.

What can Large Language Models Do?

Originally, language models were designed to handle tasks solely related to text, including:

Machine Translation: Converting text from one language to another, such as translating English to Spanish.

Sentiment Analysis: Determining whether a given piece of text expresses a positive, negative, or neutral sentiment.

Summarisation: Condensing longer texts into concise summaries while retaining the main points.

Text Generation: Creating human-like text based on prompts or inputs, useful for writing articles, stories, or dialogue.

Answering Questions: Often through a chat interface, LLMs can respond to questions either based on their pre-existing knowledge or using additional context provided by the user.

Writing Computer Code: LLMs have proven effective at generating code snippets or entire programs, aiding in software development across various applications.

However, many LLMs have now evolved into multimodal models, which means they can process and generate not just text, but also other types of data, such as images and audio. Multimodal models extend the capabilities of traditional LLMs by integrating different types of information, allowing them to perform a wider variety of tasks.

How Multimodal Models Enhance Capabilities?

AI DATA CERT

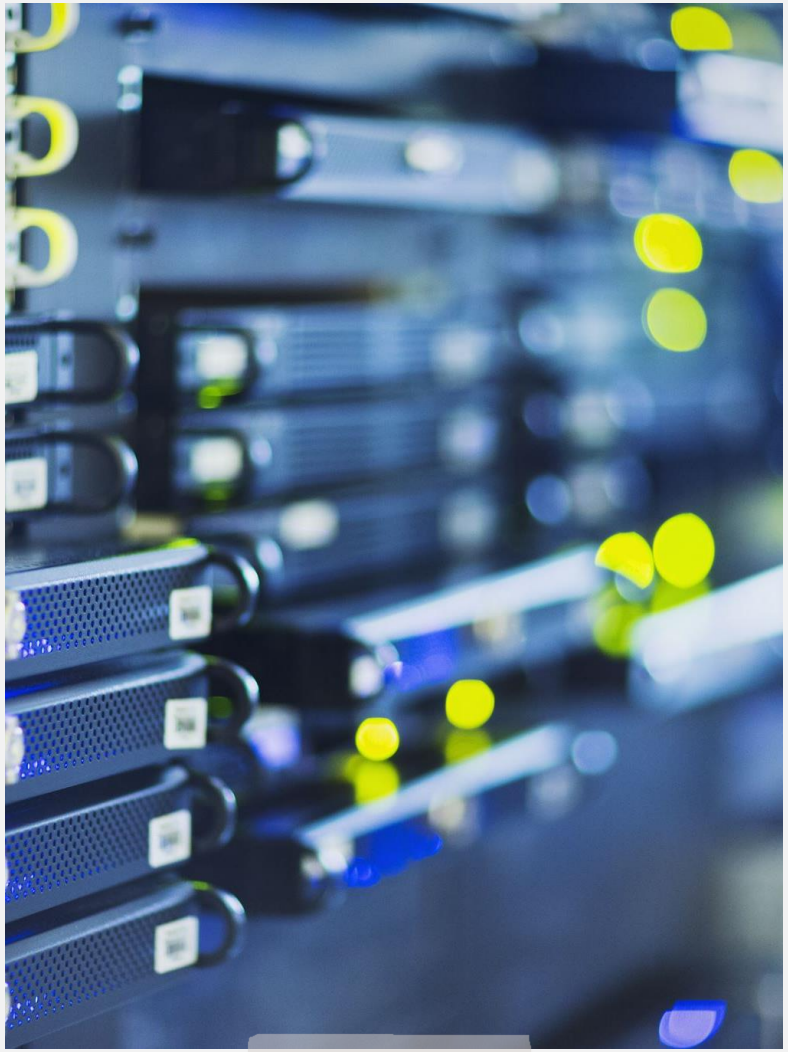
While the core architecture of these models is still focused on text, their deep understanding of language and context allows them to extend their capabilities beyond traditional text processing. Here are some of the new tasks that multimodal LLMs can handle:

Automated Image Captioning: Generating descriptive captions for images, which is particularly useful for making visual content accessible to users with visual impairments or for enhancing content management systems.
Creating Marketing Materials: Designing posters, flyers, or other marketing content by combining textual inputs with relevant images, thus streamlining the creative process.
Medical Diagnoses: Assisting in diagnosing medical conditions by analysing both textual information (like patient history) and visual data (such as X-rays or MRI scans). This integration of text and images can improve the accuracy of diagnoses and support healthcare professionals.
Speech Recognition and Generation: Converting spoken language into text and vice versa, which is essential for virtual assistants, transcription services, and accessibility features.
Visual Content Creation: Generating images based on textual descriptions, which can be applied in creative industries like advertising, design, and entertainment.
Interactive Learning Tools: Enhancing educational experiences by providing dynamic content that includes text, images, and audio to cater to different learning styles and preferences.
Multimodal Search: Allowing users to search for information using a combination of text and images, which can improve the relevance and accuracy of search results, particularly in complex or visual-heavy queries.

These examples highlight how the multimodal capabilities of modern LLMs expand their potential applications far beyond traditional language tasks, making them versatile tools in a wide range of fields.

AI DATA CERT

Challenges and Limitations



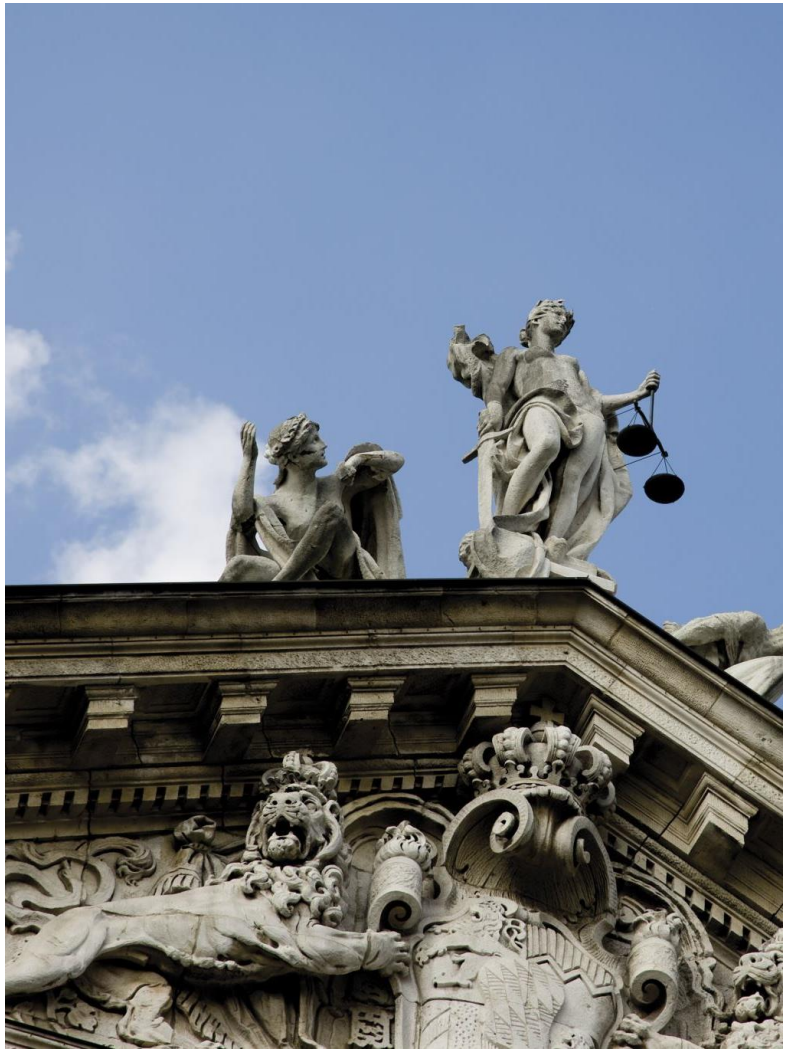
Technical Challenges

LLMs require an immense amount of computational power to train. For instance, training GPT-4 is estimated to have cost around $100 million, and the cost for future models is expected to be exponentially higher. The traditional approach to creating more intelligent models has been to increase the size of the architecture. However, advances in computing technology and the development of more efficient networks are helping to manage the complexity and scale of these models.

Another challenge is the availability of high-quality training data. LLMs have typically learned from freely available data on the internet, but there is concern that this data may be insufficient to support future models. To address this, we are seeing more partnerships between content creators and LLM developers to acquire proprietary data for training. This not only increases the volume of data but also improves its quality. Additionally, synthetic data, generated by other AI systems, is being explored as an alternative source of training data.

Compliance Challenges

When ChatGPT 3.5 was initially released, it became apparent that the model might be using user interactions (such as prompts and responses) to further train itself. This raised significant privacy concerns, particularly for organisations wary of inadvertently sharing sensitive data. Although features have been implemented to prevent models from retraining on user data, privacy remains a critical issue. The capability of these models is so advantageous that organisations often find it necessary to provide substantial information, which continues to fuel data privacy concerns. Many leading LLM providers now offer commercial licenses that include enhanced data protection and safety measures. While many companies comply with data regulations such as GDPR, it is essential to review specific agreements to ensure they meet your organisation’s requirements.



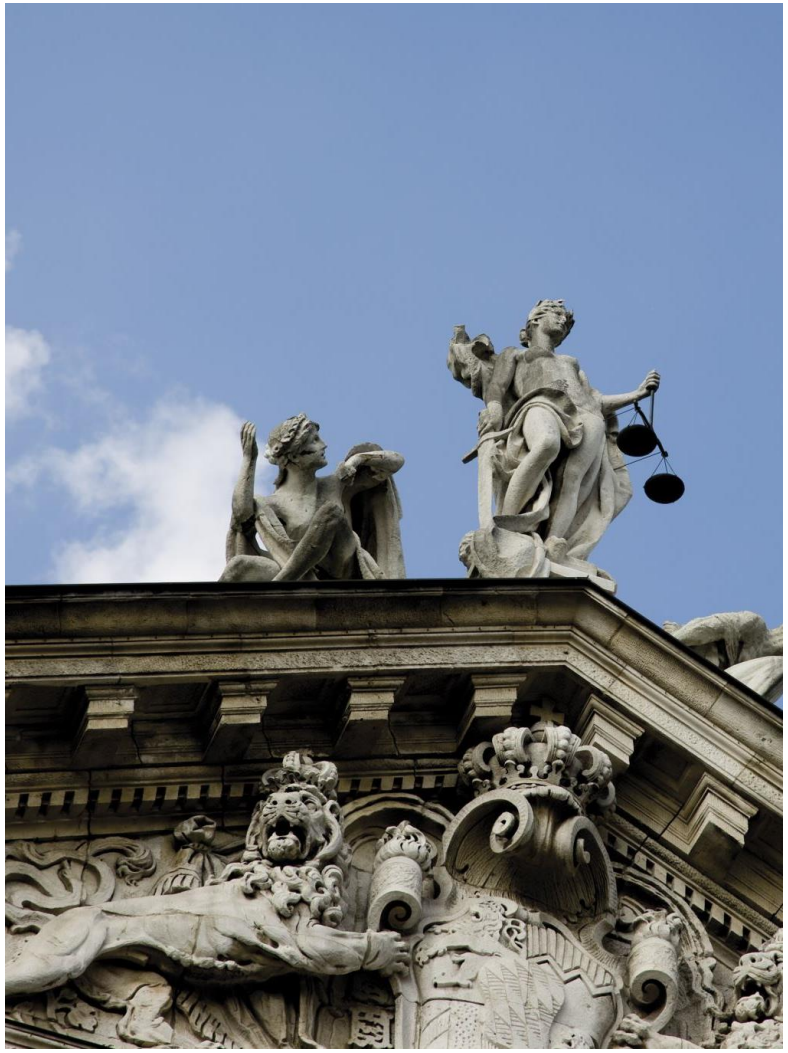
Ethical Challenges

LLMs also present numerous ethical challenges. These models are trained on vast amounts of publicly available data, which inherently contains biases. As a result, LLMs may inadvertently learn and perpetuate these biases, which can be reflected in their outputs. Although LLM providers have developed techniques to reduce bias, completely eliminating it is a formidable challenge.

Another significant limitation is the lack of explainability and transparency in LLMs. In many critical areas, such as exam grading or automated loan approvals, it is crucial to understand how decisions are made. However, given the complexity of models with hundreds of billions of parameters, providing true explainability is currently unfeasible.

The ability of LLMs to generate human-like text also raises concerns about misinformation. These models can produce vast amounts of convincingly realistic but false information, which can be used to manipulate public opinion, influence political decisions, or sway markets. With the advent of multimodal models capable of generating not just text but also images and videos, the potential for manipulation increases. While some providers have implemented restrictions to prevent the misuse of their models for these purposes, responsible use of LLMs is not uniformly enforced across the industry.

Training LLMs requires significant computational resources, leading to high energy consumption and substantial water use for cooling the data centres involved. As the world strives to decarbonise economies, the energy-intensive nature of LLMs poses a challenge. Even the inference phase—when the model is used—consumes more energy than traditional methods of information retrieval. However, some, like Sam Altman, CEO of OpenAI, argue that these new types of AI could help address the climate crisis by optimising various processes and generating innovative solutions.



Safety Challenges

The significant power of these algorithms necessitates careful consideration of the potential harm they can cause. Many model providers have implemented "guardrails" designed to prevent the generation of harmful or inappropriate content, such as blocking outputs related to dangerous or explicit topics. However, these safeguards are not foolproof, and in some cases, harmful instructions—such as those for creating dangerous substances or encouraging harmful behaviours—can still be extracted from the models. Companies like Anthropic are actively working on improving these safety mechanisms, but the challenge remains substantial.

Impact on the Workforce

Looking further into the future, the impact of LLMs on the workforce could be profound, potentially altering how society operates. These models have the potential to replace jobs across various sectors, from graphic design and advertising to executive roles within companies. If LLMs become significantly more effective than human workers, we could witness large-scale unemployment, leading to discussions about the necessity of a universal basic income.

Conclusion

In conclusion, Large Language Models represent a transformative advancement in artificial intelligence, ushering in a change comparable to the Industrial Revolution. Just as the Industrial Revolution reshaped industries and societies, LLMs have the potential to transform how we interact with technology and process information. While challenges such as significant computational demands, bias, and privacy concerns exist, the incredible capabilities of these models to understand and generate human-like text—and even integrate multiple forms of data—open up exciting possibilities across a wide range of fields.

Learn more at **www.aidatacert.com**

AI DATA CERT