

The background is a dark teal color with a faint world map. Overlaid on the map are various elements: a bar chart on the left, a flag with stars on the left, and large, semi-transparent text in various languages including 'English', 'Italiana', 'Deutsche', '日本語', and '中文'.

CHAPSVISION
DATA MAKE SENSE

WHITE PAPER

CHOOSING BETWEEN LLM AND TRANSLATION AI:

THE WINNING STRATEGY
FOR YOUR BUSINESS

INTRODUCTION CLARIFYING THE STAKES OF TECHNOLOGICAL CHOICES	5
CHAPTER 1 LLMS AND THEIR MANY APPLICATIONS	7
CHAPTER 2 LLMS: IMMENSE POTENTIAL AND CHALLENGES	9
CHAPTER 3 THE ADVANTAGES OF SPECIALIZED AI FOR TRANSLATION	12
CHAPTER 4 HOW TO CHOOSE THE RIGHT SOLUTION FOR YOUR BUSINESS?	15
CHAPTER 5 CASE STUDIES AND FEEDBACK	18
CONCLUSION	20



INTRODUCTION

CLARIFYING THE STAKES OF TECHNOLOGICAL CHOICES

CLARIFYING THE STAKES OF TECHNOLOGICAL CHOICES

In an increasingly globalized world, a company's ability to communicate effectively in multiple languages is crucial for its growth. Whether entering new markets, localizing websites, or handling legal and technical documents, machine translation has become a strategic asset.

Artificial intelligence (AI) has significantly transformed translation technologies, automating and enhancing translation quality to unprecedented levels. More recently, the rise of generative AI has opened incredible possibilities, including the ability to generate text tailored to different contexts. However, despite the promise of this technological revolution, companies must thoroughly understand the uses and limitations of these tools. Relying solely on the raw power of models is insufficient; it is essential to adapt the technology to specific needs to avoid errors, inaccuracies, and unnecessary costs.



Today, two approaches stand out in this field: large language models (LLMs) and AI specialized in translation. While LLMs offer impressive flexibility, specialized AI focuses on delivering high-precision performance in specific contexts.

This white paper aims to provide an in-depth analysis of the advantages and limitations of these two technologies. The goal is to help businesses make informed choices based on their machine translation needs, considering factors such as accuracy, data security, and adaptability to specialized environments.



CHAPTER 1

**LLMS AND
THEIR MANY APPLICATIONS**

WHAT IS AN LLM?

A large language model (LLM) is an artificial intelligence trained on vast volumes of textual data using Transformer architecture (or “self-attentive” models). These models are trained on billions, or even trillions, of words, enabling them to learn complex relationships between words and generate text that appears natural and fluent. What sets LLMs apart is their versatility: they can be used for a variety of linguistic tasks, from content generation to machine translation. matter knowledge.

The most well-known example is GPT4. These models can handle multiple languages and generate text in several languages with a degree of coherence based on the context provided. However, despite their impressive flexibility, LLMs face limitations when dealing with specialized domains requiring precise terminology and deep subject matter knowledge.



DIVERSE APPLICATIONS OF LLM

LLMs are not limited to machine translation. Here are some practical applications across different sectors:

- **Content Generation:** LLMs are used to write blog posts, corporate reports, video scripts, and even books. For example, marketing firms automate the creation of newsletters tailored to customer preferences.
- **Writing Assistance:** Tools like Grammarly leverage LLMs to correct grammar and suggest clearer, more impactful rephrasing.
- **Assisted Programming:** Models like GitHub Copilot generate lines of code based on the task at hand, helping developers solve technical problems faster.
- **Sentiment Analysis:** LLMs are used to process and analyze large datasets, such as customer satisfaction surveys or social media comments.
- **Chatbots and Customer Service:** Companies like Sephora and H&M use LLM-based conversational agents to interact with customers, automatically answering common questions and resolving minor issues.

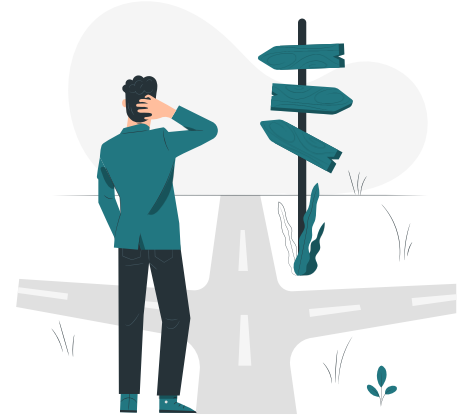
CHAPTER 2

LLMS: IMMENSE POTENTIAL AND CHALLENGES

THE "HALLUCINATION" PHENOMENON IN LLMS

Hallucinations are an inherent problem in LLMs. These models do not truly understand the world they describe but rather produce word sequences based on probabilities. When LLMs encounter incomplete or ambiguous information, they may generate responses that appear coherent but are actually incorrect or fabricated.

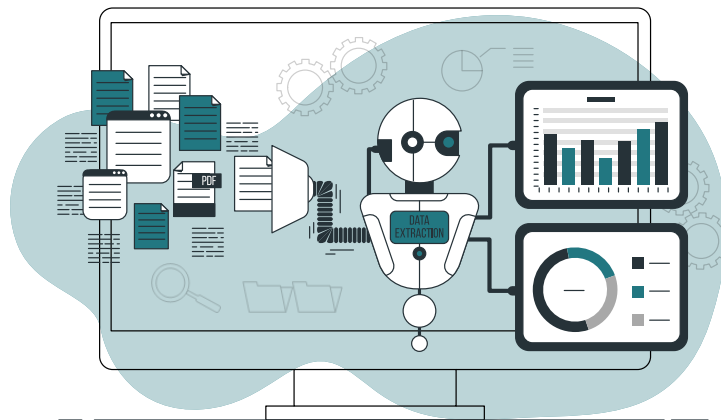
For instance, a law firm using an LLM to translate a complex contract could face severe legal consequences if the model, lacking sufficient training in legal terminology, produces inaccurate translations.



LLM TRAINING: MASSIVE DATA AND RESOURCES

Training LLMs requires huge amounts of data and computing resources. To achieve a sufficient level of performance, a model must be trained on huge text corpora from varied sources, such as books, scientific articles, web pages, and online forums.

For example, training GPT-4 required weeks of computation on supercomputers equipped with thousands of processors, representing a significant financial and energy cost. Increasing concerns have arisen about the environmental impact of these models, given their massive energy consumption.



LIMITS AND ETHICAL CHALLENGES

Beyond technical challenges, LLMs raise ethical issues. These models are often trained on publicly available textual data, which may contain social, cultural, or linguistic biases. For example, if a model's training data includes gender stereotypes, it risks reproducing them in its responses.

Data confidentiality is also a major concern. LLMs may be trained on sensitive or private information, posing risks to compliance with regulations like GDPR. Companies must ensure that the data used to train and run these models does not compromise the security of confidential information.



Lastly, the issue of intellectual property arises, particularly when LLMs are trained on copyrighted works such as books or scientific articles, raising legal questions about the permissible use of these materials for AI training.

CHAPTER 3

THE ADVANTAGES OF SPECIALIZED AI FOR TRANSLATION

MACHINE TRANSLATION: LLM VS. SPECIALIZED AI

While LLMs can translate text from one language to another, their flexibility can become a weakness when translating in highly specialized fields. For example, in domains like law or medicine, precise terminology is crucial. A generalist model like an LLM might misinterpret specific terms or omit important nuances.

Specialized AI, on the other hand, is designed for extreme precision in specific contexts. Trained on specialized corpora (legal, medical, technical), it handles complex translations with much higher accuracy than generalist LLMs. Furthermore, a translation-specialized AI has been specifically trained for the task of converting text from a source language to a target language. Translation models are thus less prone to hallucinations since they are trained to transform text rather than create it.

WHY CHOOSE A SPECIALIZED AI?

Companies operating in regulated or technical sectors must choose a machine translation solution that guarantees maximum accuracy. Specialized AIs offer this advantage through their targeted training. They incorporate precise terminologies and can understand the specific context of a text, significantly reducing the risk of translation errors.

Specialized AIs can also be customized to meet a company's unique needs, including the addition of specific glossaries, training on internal corpora, and adapting to cultural nuances of each target language.



DATA SECURITY AND COMPLIANCE

Another advantage of specialized AI is its ability to meet data security and regulatory compliance requirements. Unlike generalist LLMs, which often rely on public data, specialized AIs can be trained internally on confidential data with high-security guarantees. This ensures the confidentiality of sensitive information, particularly in sectors like healthcare or financial services.

COMPLEMENTARITY BETWEEN LLMS AND SPECIALIZED AI FOR TRANSLATION

It is essential to note that generative AI, LLMs, and specialized translation AI share a common technological foundation: transformer architecture. However, these models have been trained for distinct uses. Generative AI can produce varied content, such as images, videos, sound, code, and 3D models. LLMs focus on text generation with applications such as comprehension, analysis, content creation, or chatbots. Lastly, specialized translation AI focuses exclusively on the precise conversion of text from a source language to a target language.

Despite this specialization, the common technology allows for functional complementarities between these models—some already well established, others yet to be explored. For example, LLMs can be used to enhance the quality of translation AI training data by classifying, filtering, and cleaning linguistic corpora before training.

A hybrid approach combining LLMs and specialized AI is also possible.

Take, for instance, conversational LLMs, often optimized for English. In this case, a specialized translation AI can play a crucial role by translating initial queries into English for the chatbot, then relaying responses in the user's native language, thereby improving the multilingual user experience.

Finally, LLMs, with their ability to creatively rephrase and simplify texts, can complement translation AI by refining translations, reformulating or simplifying while preserving the original meaning.



CHAPTER 4

HOW TO CHOOSE THE RIGHT SOLUTION FOR YOUR BUSINESS?

MACHINE TRANSLATION: LLM VS. SPECIALIZED AI

For a company, choosing between an LLM and specialized AI for machine translation depends on several factors. Here are some criteria to consider:

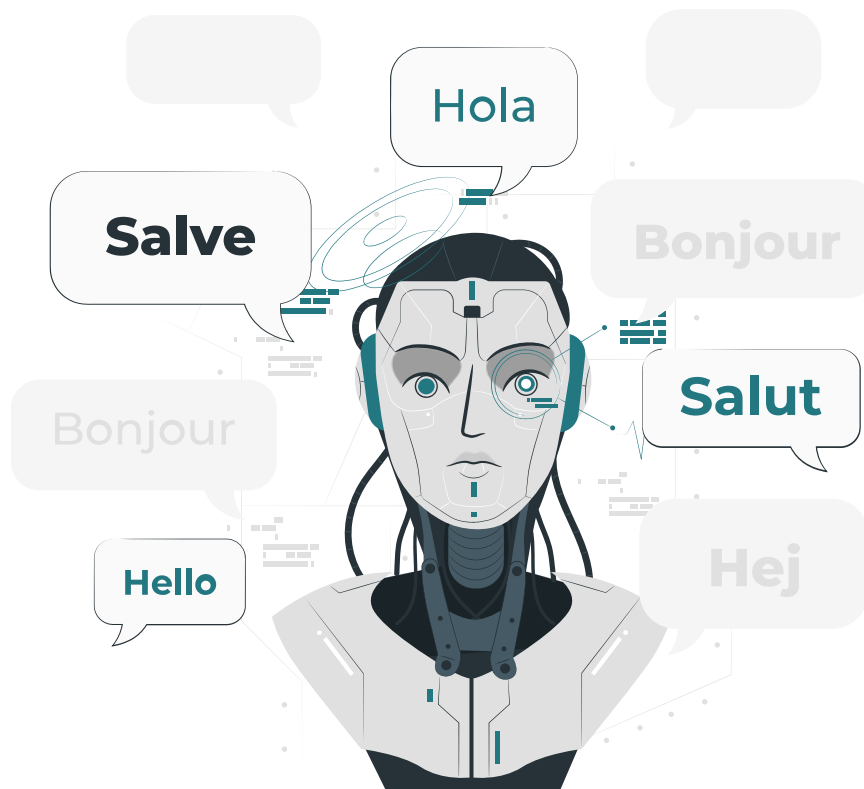
- ❖ **Translation Accuracy:** If your sector requires precise translations (law, healthcare, finance, sensitive industries), the consequences of incorrect translations can be severe. It is crucial to assess the level of accuracy needed in the translated texts, especially in technical and regulated fields.
- ❖ **Data Confidentiality:** Do your documents contain sensitive or confidential information? Ensure that the chosen solution offers robust guarantees regarding data security and confidentiality, particularly in compliance with regulations such as GDPR.
- ❖ **Budget and Resources:** Some translation solutions may require significant hardware resources and incur high operational costs. If your budget is limited, it's essential to analyze long-term costs, especially if you have large volumes of text to translate regularly.
- ❖ **Deployment:** Does your solution need to be flexible and integrate easily into your existing systems (cloud, on-premise, API)? Check the available deployment options and ensure that the solution can be adapted to your current technical infrastructure.
- ❖ **Performance:** Translation speed is a key factor to consider. If you need real-time translations or require high responsiveness, it is important to ensure that the solution can deliver translations quickly, even for large volumes. Evaluate response times and processing capabilities according to your operational needs.
- ❖ **Operating Costs:** Some solutions may be costly due to high computational and infrastructure needs. This can be a critical factor if you handle large volumes of text regularly. Compare operating costs and ensure the solution is cost-effective in the long run.
- ❖ **Tools and Integrations:** Does the translation solution need to integrate easily with your existing professional tools, such as CMS platforms, CRM systems, or other specialized software? Check whether it offers compatible connectors or APIs for quick and efficient integration into your workflows.
- ❖ **User Management:** If you require simplified user and access management, make sure the solution offers intuitive interfaces and easy administration. User management should be adaptable to your needs for control and monitoring of usage.

MACHINE TRANSLATION: LLM VS. SPECIALIZED AI

In summary, choosing the right machine translation solution for your business requires a precise evaluation of your needs and constraints. Specialized solutions often offer advantages in terms of accuracy, speed, and seamless integration into complex technical or regulatory environments. They are optimized to efficiently handle large volumes of data while ensuring enhanced security, particularly for sensitive documents.

On the other hand, LLMs have some limitations for specialized applications. While versatile and capable of handling various linguistic tasks, they may lack accuracy in specific technical fields and require significant resources, which can increase operational costs. Moreover, their response time may be slower when processing large volumes of data or complex content, and their data confidentiality may present higher risks due to the nature of the data used for training.

Ultimately, if your company operates in a sector where precision, security, and performance are crucial, an AI specialized in translation is often a more suitable choice. It meets high demands while offering a more cost-effective solution tailored to your specific needs.



CHAPTER 5

CASE STUDIES AND FEEDBACK

CASE 1: MANAGING A GLOBAL E-COMMERCE PLATFORM WITH HIGHLY CREATIVE CONTENT

With a large internal localization team—no fewer than 100 people—and an increasing volume of translation for each market (about 20 million words translated per year), the European leader in sporting goods needed a solution that was easy to use and fast to execute. Even more so for creating highly creative content (marketing, communication, and e-commerce), which requires a highly qualified linguistic team. The specialized machine translation solution implemented allowed them to offload the problem of additional content translations and deliver 25% more multilingual content with the same team and the same level of quality.

CASE 2: ADAPTING MACHINE TRANSLATION TO INDUSTRY JARGON

A language service provider that is part of one of the largest networks of linguistic experts (active in no fewer than 60 countries and more than 100 languages).

Every day, a translation agency helps its clients with their international communication by offering translation and localization services. Among its clients are several industries (construction, civil engineering, energy, etc.) that require a perfect mastery of complex lexicons and sometimes challenging subjects. The specialized machine translation technology allows them to train their own translation models for specific domains or clients, enabling greater precision and delivering high-quality content in a short time.

CASE 3: SECURING THE TRANSLATION PROCESS

A Swiss bank chose to provide all its employees with a specialized and secure machine translation solution to meet the Group's security requirements by controlling the processing of translated data, significantly reducing the risk of confidential data leaks.

To discover more use cases and benefits of specialized AI in machine translation, feel free to consult our study:



**ROI of Machine Translation
Overview of User Feedback.**

CONCLUSION

In the context of increasing internationalization, machine translation is becoming a strategic lever for companies. The choice between large language models (LLMs) and specialized AI depends on specific criteria such as accuracy, data security, and technical or sector-specific needs.

LLMs stand out for their versatility and ability to adapt to multiple linguistic tasks, ranging from translation to customer support. However, their flexibility can sometimes be a disadvantage when faced with precision requirements in specialized fields. Specialized AIs, on the other hand, offer deep expertise in regulated or technical sectors, ensuring high-quality results and reducing the risk of errors.

For companies dealing with large volumes of translation or sensitive documents, a specialized AI solution is often more cost-effective and better suited in terms of compliance and security. Nonetheless, a hybrid approach combining the flexibility of LLMs and the precision of specialized AI can also represent an effective compromise, depending on the needs of each organization.



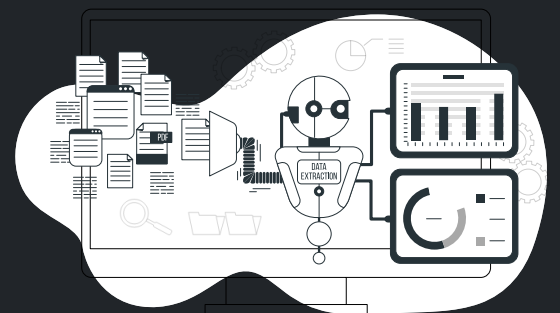
ABOUT CHAPSVISION

As a specialist in sovereign data processing, ChapsVision enables businesses and public authorities to succeed in their digital transformation and create value thanks to a software suite based around its system for exploiting massive and heterogeneous data.

Through substantial R&D investment in the processing of massive, heterogeneous data, complemented by a sustained strategy of targeted acquisitions and international expansion, ChapsVision has rapidly built up a coherent group comprising two divisions addressing sectors with a strong data footprint. One is dedicated to customer engagement, unified commerce and economic and strategic intelligence for businesses, while the other offers sovereign solutions dedicated to cyber intelligence and cybersecurity for the defence, intelligence and security markets.

AI CAPABILITIES IN THE CHAPSVISION GROUP

Artificial intelligence (AI) plays a crucial role in the various activities of the ChapsVision group, with extensive applications offering innovative and effective solutions for data analysis and processing. In the field of audio, AI enables speech-to-text transcription, noise suppression, and the detection of emotions, events and speakers. For text, it facilitates natural language processing (NLP), large language models (LLM), sentiment analysis, entity and relationship detection, as well as summarisation and translation. For images, AI detects specific objects and enables text to be read. For videos, it detects scenes and objects, and for documents, it extracts tables and images.



DATA AND AI: CHAPSVISION TOOLS

The ChapsVision group offers advanced tools, enhanced with artificial intelligence, that enable you to exploit your customer data to optimise your customer relations. ChapsVision is renowned for its Customer and Sales Force CRM, Marketing Automation, Unified Commerce, Market Intelligence, Single Customer Repository (SCR), Customer Data Platform (CDP) and Data Mining solutions. Thanks to our innovative solutions, our customers benefit from a 360° view of their business ecosystem and customer knowledge, enabling them to make decisions quickly, optimise their revenues and customer loyalty, while reducing risks in a complex and increasingly competitive environment.





CHAPSVISION

DATA MAKE SENSE

**YOU WOULD LIKE TO BE PUT IN TOUCH
WITH OUR EXPERTS?**

**contact@chapsvision.com
www.chapsvision.com**