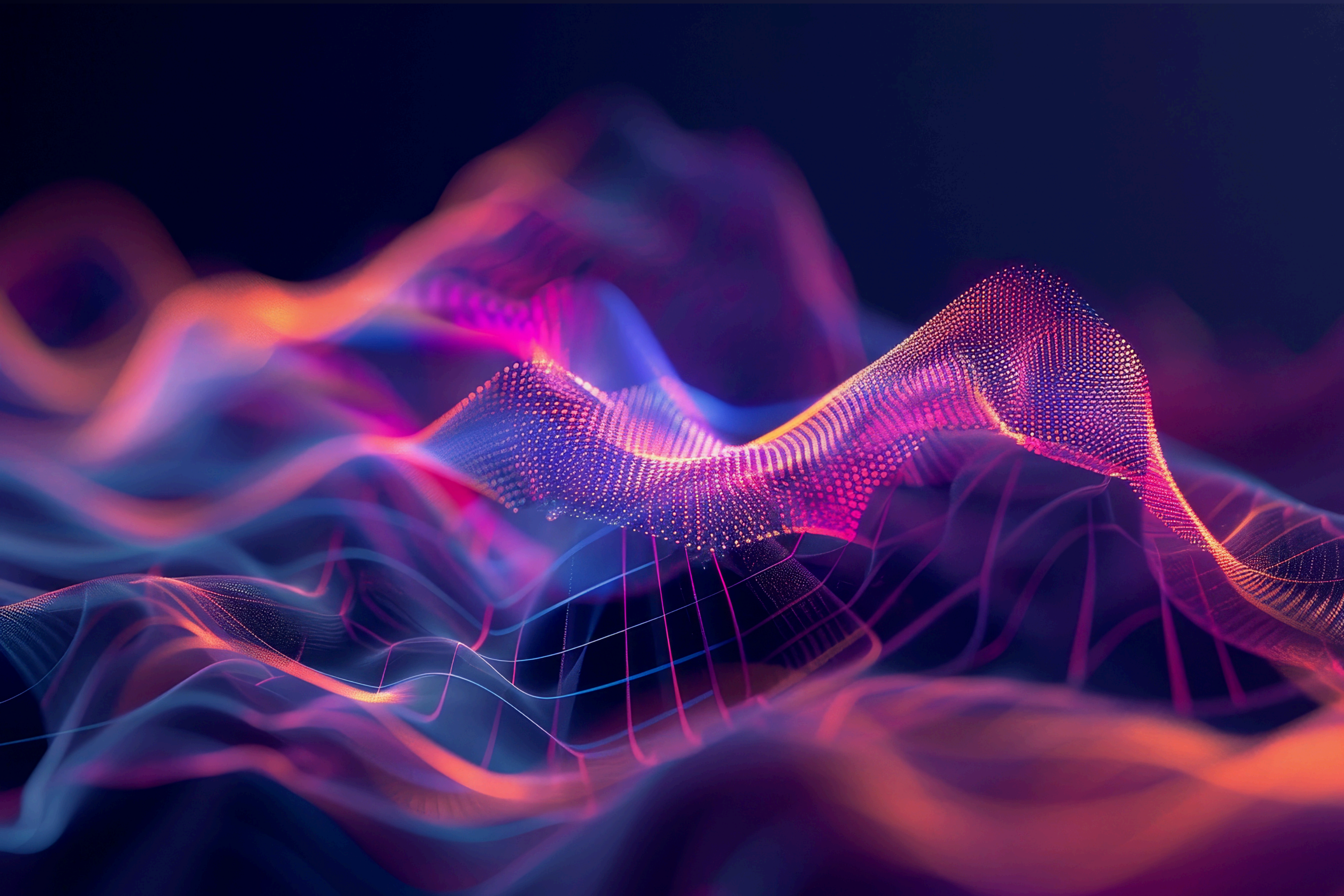


LLM GROUNDING

How DATAmundi helps
companies ground their
Large Language Models:
A Deep Dive



LLM GROUNDING

What is grounding?

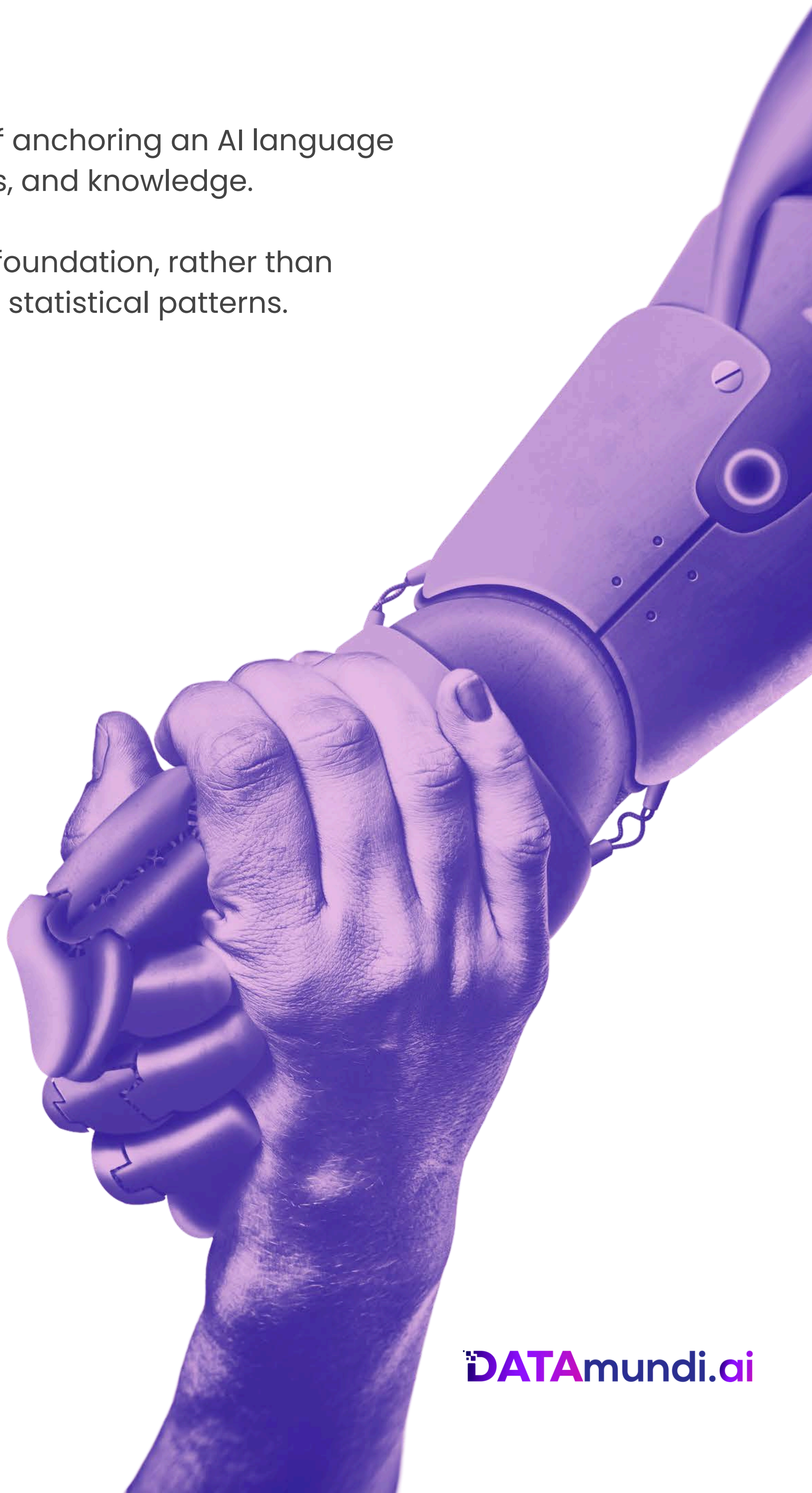
Grounding is the critical process of anchoring an AI language model to real-world facts, contexts, and knowledge.

Think of it as giving your AI a solid foundation, rather than letting it operate in a world of pure statistical patterns.

Why Grounding AI Models is Essential for Accuracy and Reliability?

Without proper grounding, language models can generate fluent but factually incorrect or contextually inappropriate responses – like a politician who confidently states that "windmills cause cancer".

As companies deploy AI in high-stakes environments like healthcare, finance, and customer service, ensuring models are well-grounded isn't just a technical necessity – it's a business imperative, and a social responsibility.



ADDRESSING THE CHALLENGES OF LARGE LANGUAGE MODELS

The rise of large language models has created new challenges for AI companies seeking to make their models more accurate and reliable. That's where DATAmundi comes in. As a leading AI Data and Language Service provider, we specialize in helping companies improve their LLMs through expert data curation and linguistic expertise.

Our journey with customers typically starts with intensive domain-specific data collection.

Whether you're building a healthcare chatbot or a financial services AI, we gather precisely what your model needs - from medical terminology databases to real-world examples of financial conversations across different markets and languages.

But collecting data is just the beginning

Our team of linguistic experts enhances your training data through professional translation, localization, and rigorous quality assurance.

We don't just translate words; we ensure that every piece of data carries **the right context and meaning for your specific use case.**

BEYOND WORDS: ENSURING CONTEXT AND CULTURAL RELEVANCE IN AI

Since the arrival of LLMs, our data team has been involved with the client science teams in projects and experiments involving:

- **Scoring, labeling, or ranking** various LLM model responses are often considered key indicators like correctness, faithfulness, and usefulness for the end user. The domains varied from generic newsworthy ones to very specialized ones like Finance, Technical, Health, IT, etc.
- **Analyzing, evaluating, and labeling** the sources of truth (evidence) the LLMs should use, sometimes also extracting from the relevant ones the paragraphs that help answer the queries
- **Labeling projects** used for detecting the sources that are trustworthy versus those that are not to be used, be it biased, offensive, or simply not useful to answer the queries
- **Creating golden answers**, questions and answers (Q&As) based on text, images, documents, tables, databases



We take bias detection seriously

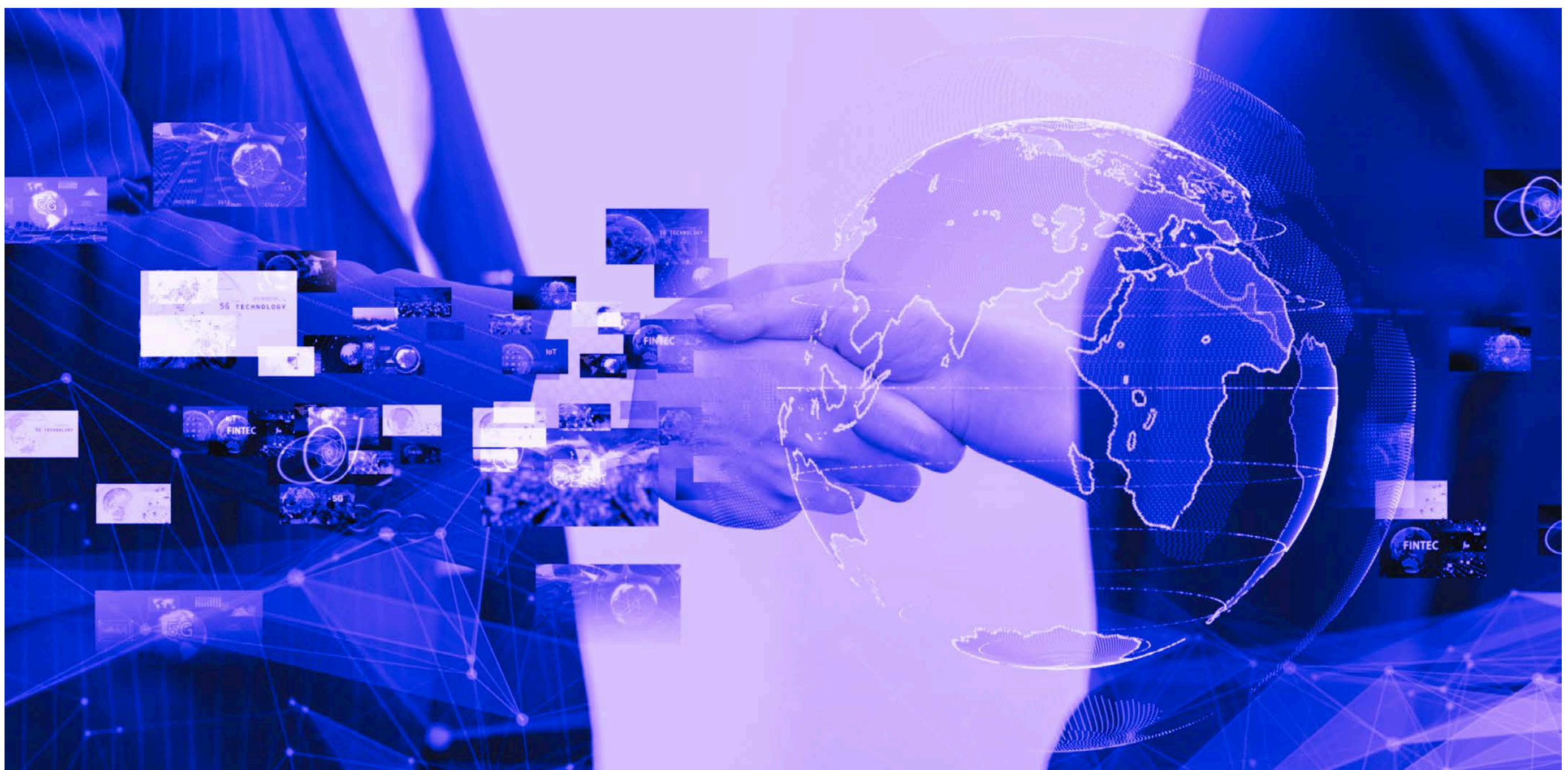
Our diverse team of experts works tirelessly to identify and mitigate cultural and linguistic biases in your training data. We ensure balanced representation while flagging potentially problematic content before it affects your model's performance.

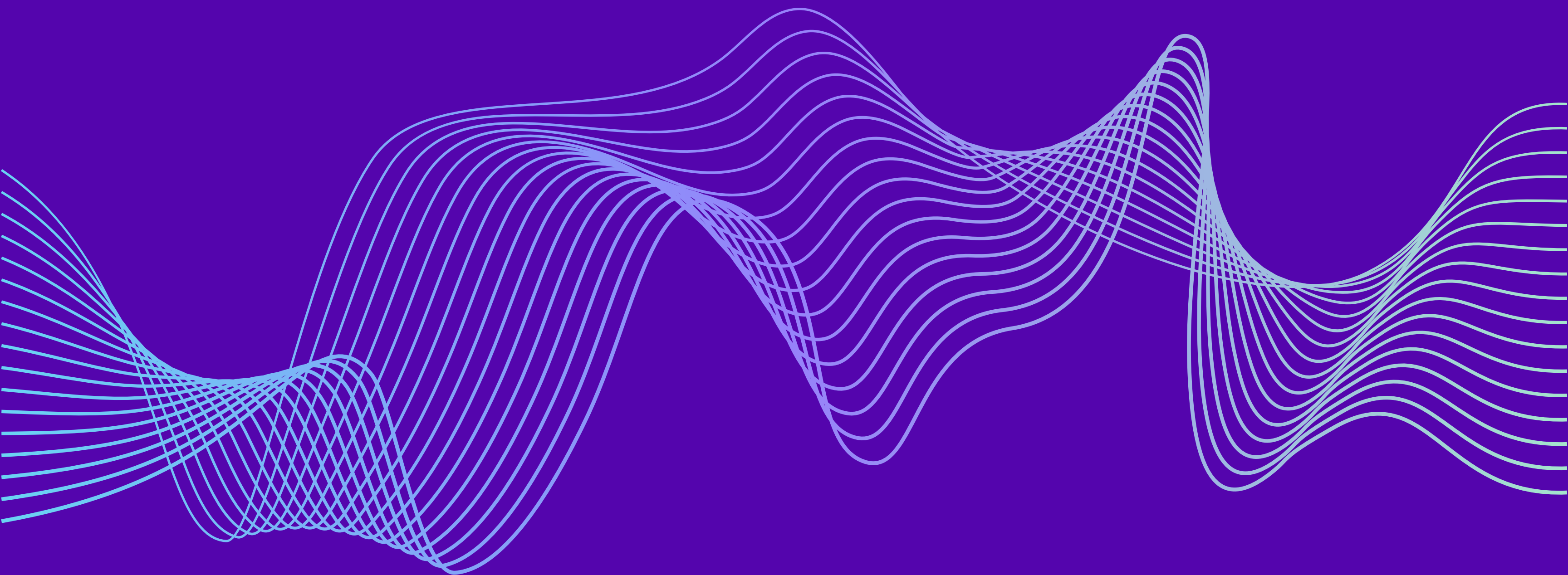
At DATAmundi, we know that better language models start with better data.

By partnering with us, you're not just getting a service provider – you're getting a team dedicated to helping your AI understand the world as it really is, in all its complexity and nuance.

Ready to take your LLM to the next level?

Let's talk about how DATAmundi can help ground your model in reality.





DATAmundi.ai

CONTACT

*A Multilingual Data
Company with Offices
Worldwide.*

www.datamundi.ai
contact@datamundi.ai