Omniscien™ TECHNOLOGIES

language STUDIO

Dion Wiggins
Founder & Chief Technology Officer,
Omniscien Technologies

Professor Philipp Koehn
Chief Scientist
Omniscien Technologies

Professor of Computer Science,
Johns Hopkins University

Reality Check: AI and Language Processing in 2025

What Works, What Fails, and What Matters Next

Agenda

This webinar will cover:

- Top 5 Advances in AI that You Will Want to Know About
- Top 5 AI Falsehoods that Most People Believe In
- AI Snapshot for the Language Services Industry
- An Introduction to People-First AI
- Why Every Organization Should Care About Digital Sovereignty
- Official Launch of Language Studio 7

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

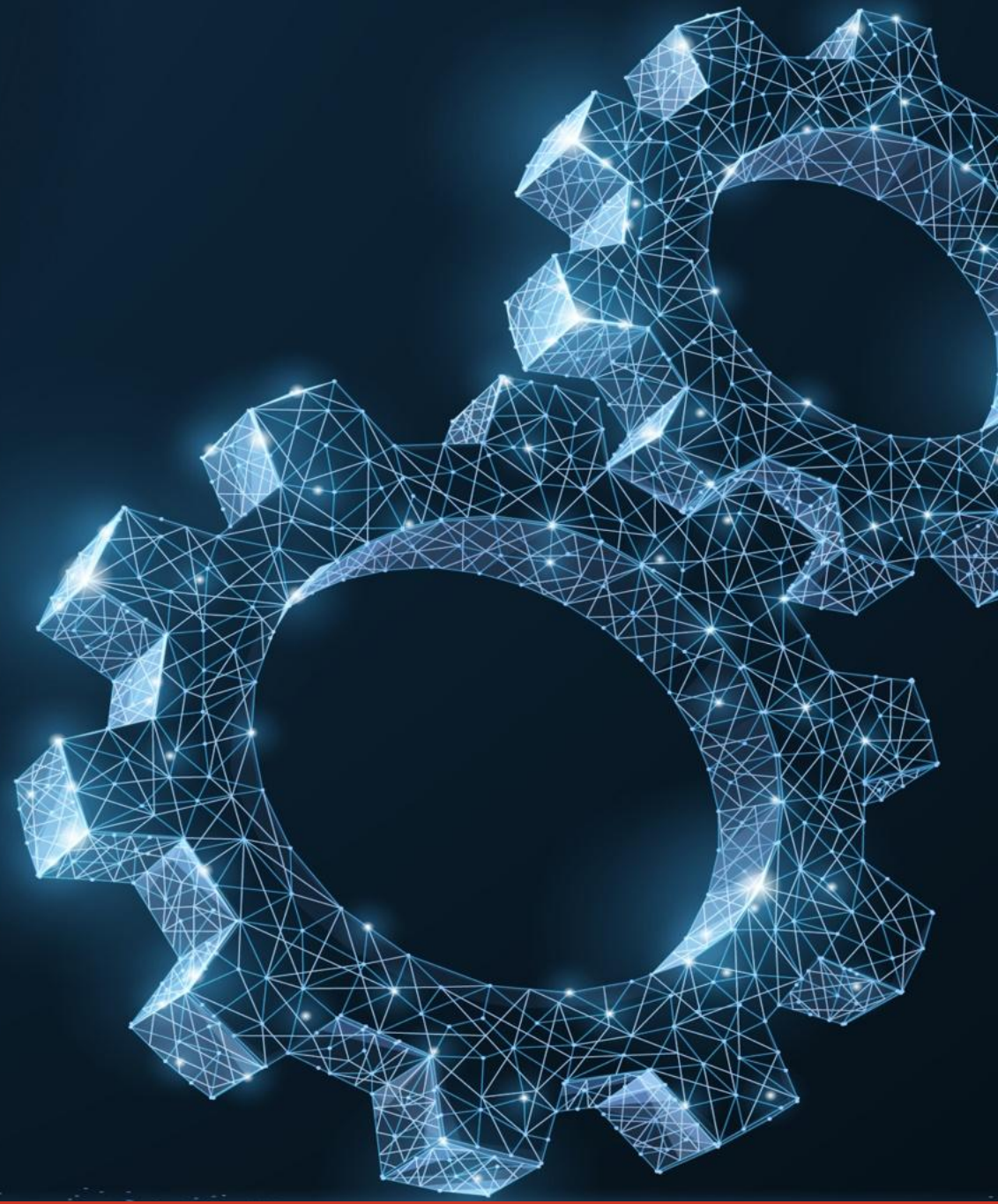
Top 5 Advances in AI that You Will Want to Know About

Omniscien™ TECHNOLOGIES

ADVANCE: #1

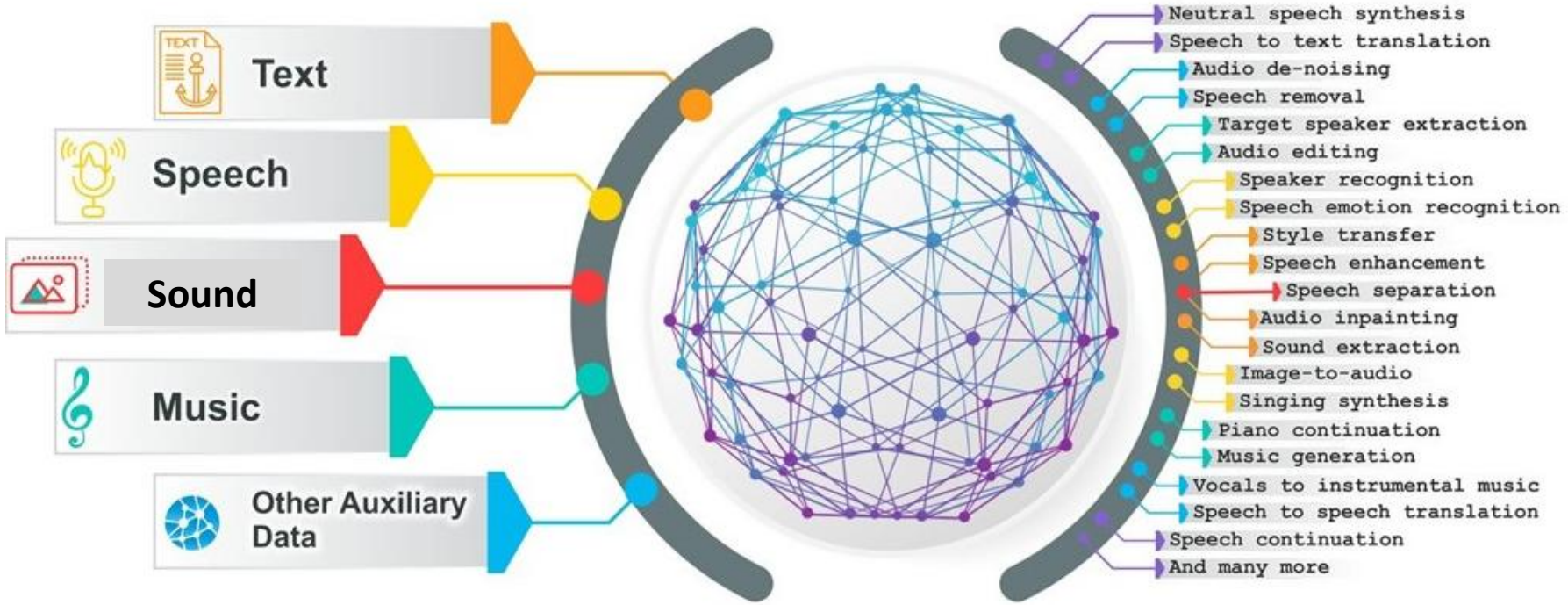
True Multimodal Reasoning

Not Just Modal Inputs



Audio Language Models

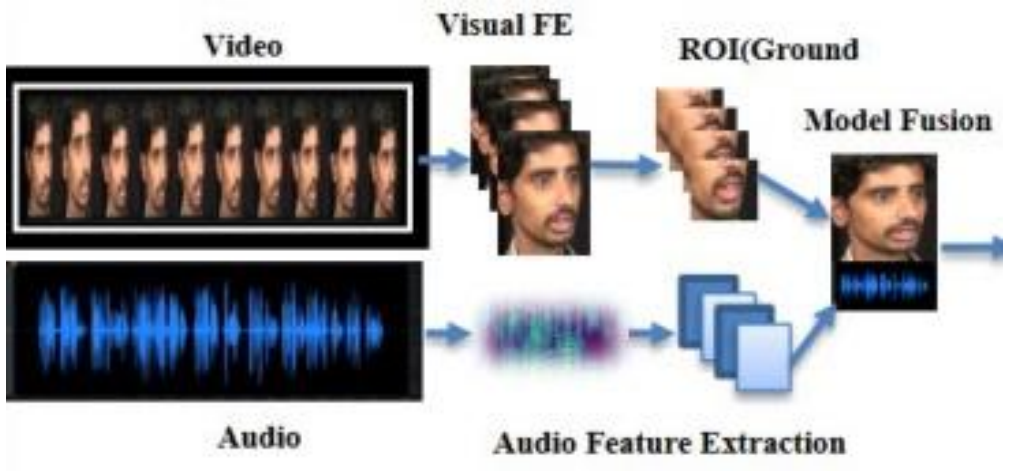
- Extension of Text LLMs with Audio

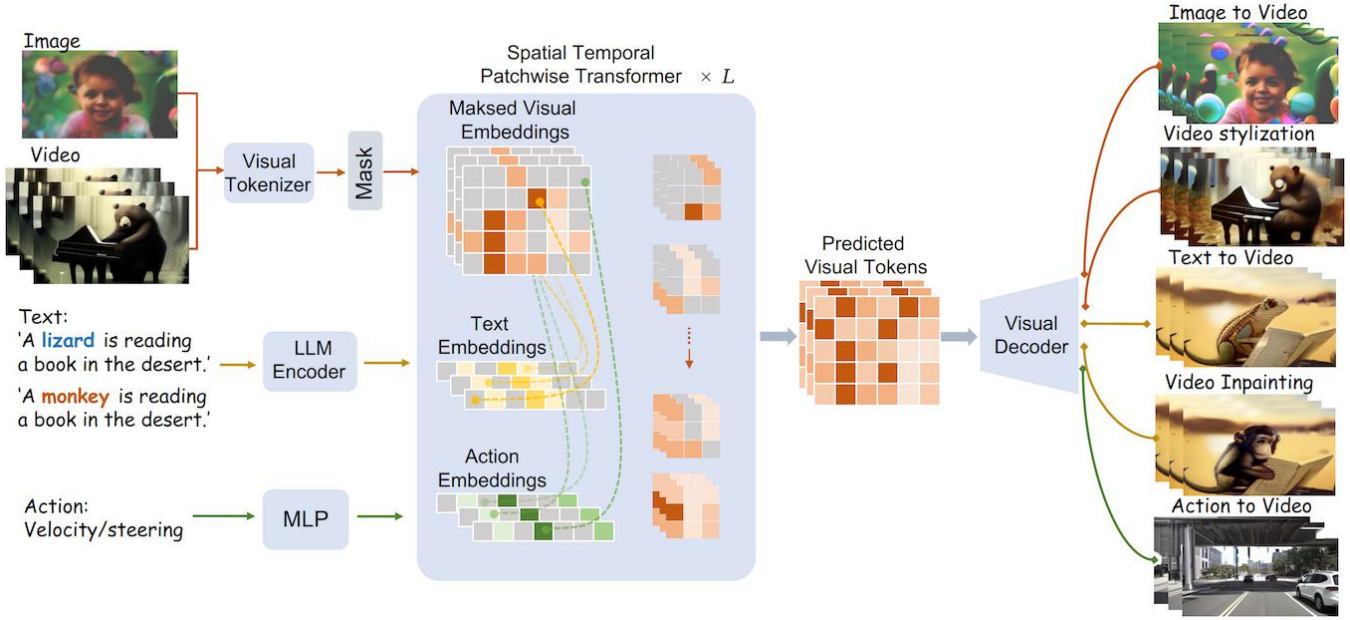


- Audio is not just speech: music, environmental sounds, etc.

Audio-Visual Language Models

Additional Extension with Image and Video



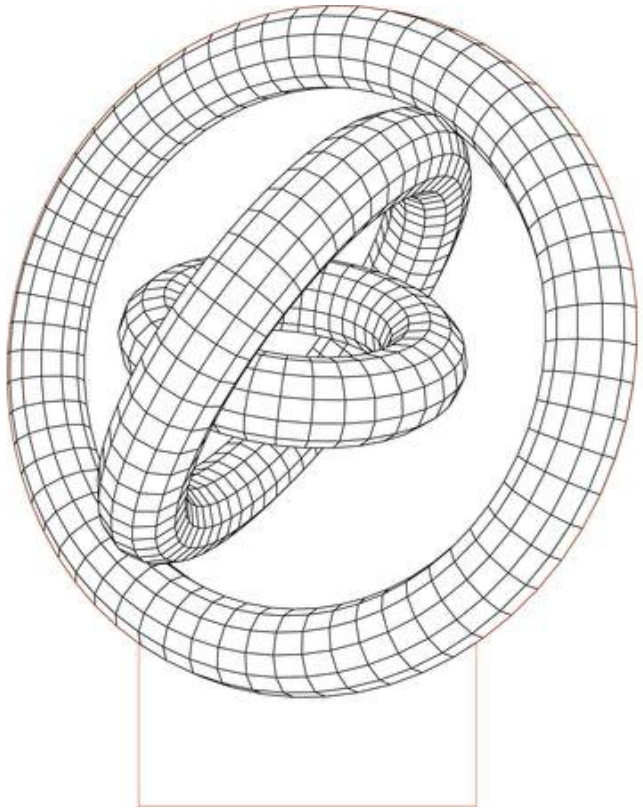


Facial expression and gestures to extract communicative information

Video/text alignment to obtain world models

Physics from Video?

Traditionally: Abstract 3D models and physical simulation



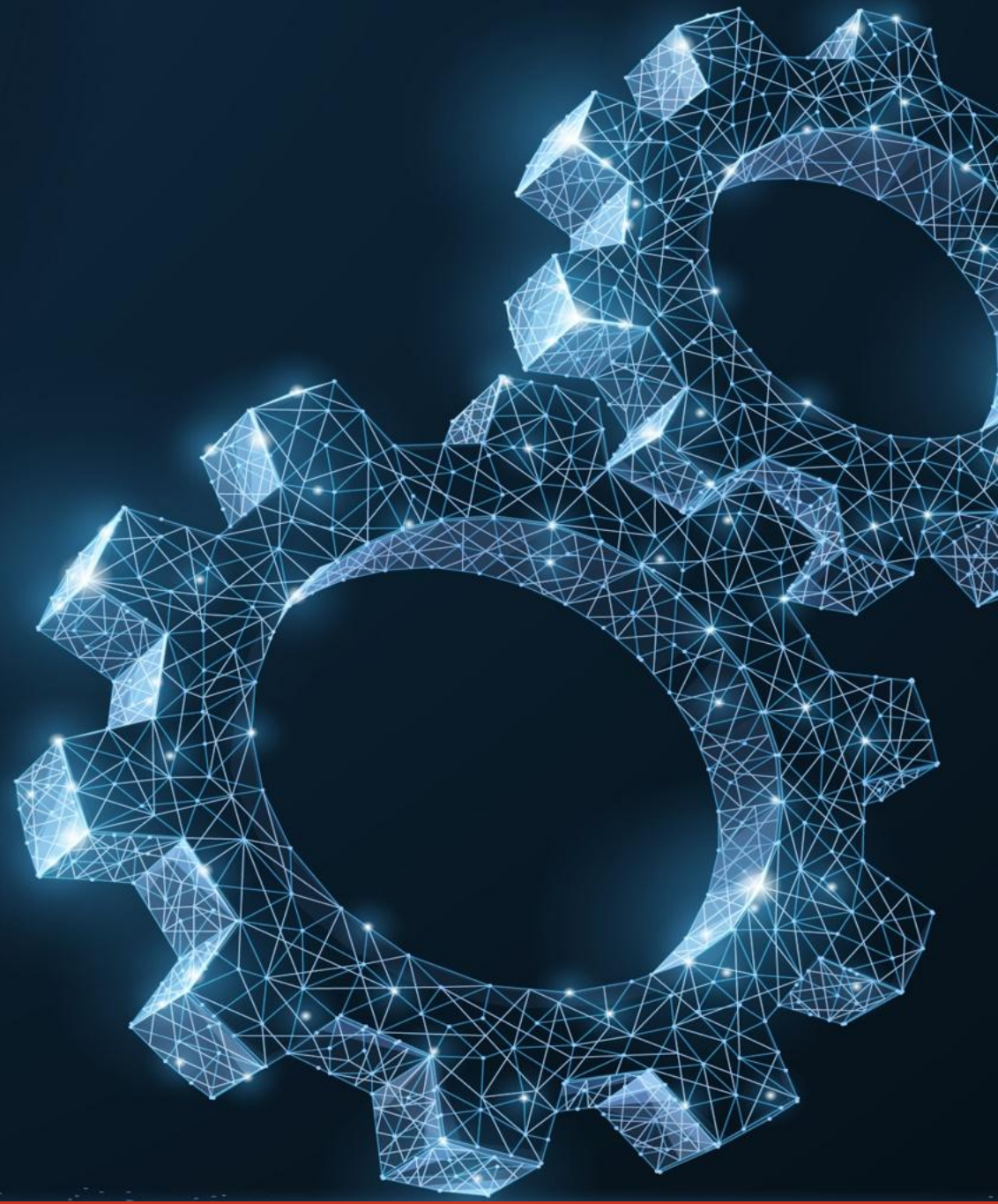
Now: learn physics from video

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

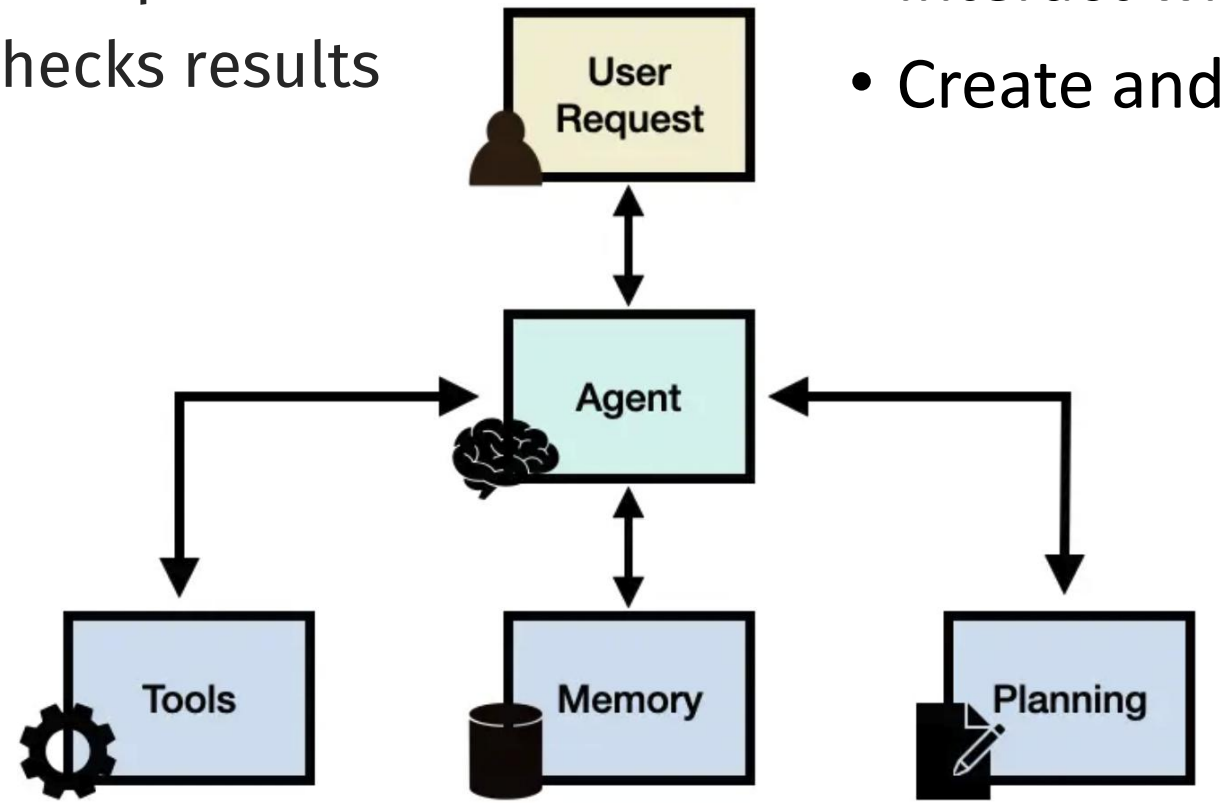
ADVANCE: #2

Agentic AI
Autonomous Planning + Tool Use



Agents

- Create plans
- Revise plan
- Calls multiple components
- Synthesizes, checks results



- Take actions, such as filling out web forms
- Interact with databases
- Create and execute code

Google (2024): Four Stages Following Professional Translation Process

Pre-drafting Research

You will be asked to translate a piece of text from English into Chinese following the five stages of the translation process. Here is the context in which the text appears:

Context: However, he said the most "responsible way" to cut spending would be to pass all 12 bills. With many Republicans reticent to see a shutdown and a 1 percent cut to defense spending, the urgency to pass the bills could see enough Republicans ally with Democrats to extend the deadline on the four spending bills and reach an agreement to fund the government, even if doing so puts Johnson in hot water with some of those on the hard right.

To start, let's do some pre-drafting research on the above context:

Research:

During this phase, thorough research is essential to address components of the context text that pose translation challenges. The goal is to establish a comprehensive translation plan that covers the following category:

- Idiomatic Expressions:
  - Identify idiomatic expressions that cannot be directly translated word-for-word into Chinese.

Idiomatic Expressions in the Context:

Here are some idiomatic expressions in the context that require careful consideration for translation into Chinese:

- **"in hot water":** This idiom means to be in trouble or facing criticism. A direct translation wouldn't convey the meaning.
- **"hard right":** This refers to the most conservative faction within a political party. A literal translation wouldn't be accurate.
- **"reticent to see":** This implies reluctance or unwillingness. A more nuanced translation is needed.

Drafting

Now, let's move on to the drafting stage.

Draft Translation:

In this phase, your primary objective is to create a draft translation that accurately conveys the meaning of the source text presented below. At this stage, it is crucial to focus on adequacy, ensuring that your translation closely adheres to the source text. Your response should conclude with the draft translation. If context is missing, generate a general translation that is adaptable to various contexts. Avoid adding any additional information not present in the source text. All elements of the source text should be present in the translation.

Give your best one translation for the following piece of text based on the pre-drafting analysis without providing alternatives:

English: However, he said the most "responsible way" to cut spending would be to pass all 12 bills. With many Republicans reticent to see a shutdown and a 1 percent cut to defense spending, the urgency to pass the bills could see enough Republicans ally with Democrats to extend the deadline on the four...

然而，他表示，削减开支最“负责任的方式”是通过所有 12 项法案。由于许多共和党人不愿看到政府停摆和国防开支削减 1%，因此通过这些法案的紧迫性可能会促使足够多的共和党人与民主党人结盟，延长四项支出法案的截止日期，并达成一项为政府提供资金的协议，即使这样做会让约翰逊与一些极右翼人士产生摩擦。

Refinement

Now let's move to the next stage.

Post-editing with local refinement

In this stage, the primary aim is to refine the draft translation by making micro-level improvements that improve the draft's fluency.

Provide only one refined translation and do not output anything else after that.

然而，他表示，削减开支最“负责任的方式”是通过全部 12 项法案。鉴于许多共和党人不愿看到政府停摆和国防开支削减 1%，通过法案的紧迫性可能会促使足够多的共和党人与民主党人合作，延长四项支出法案的截止日期，并达成一项为政府提供资金的协议，即使这样做会让约翰逊与一些强硬的右翼人士产生矛盾。

Proofreading

You are tasked with proofreading a translation that has been revised for improved fluency. The refined translation has been generated by editing the draft translation.

Proofreading and Final Editing

The goal is to provide a polished final translation of the source text. For you reference, below are the source text, the draft, and refined translations.

Source Text
However, he said t....

Draft Translation
然而，他表示....

Refined Translation
然而，他表示....

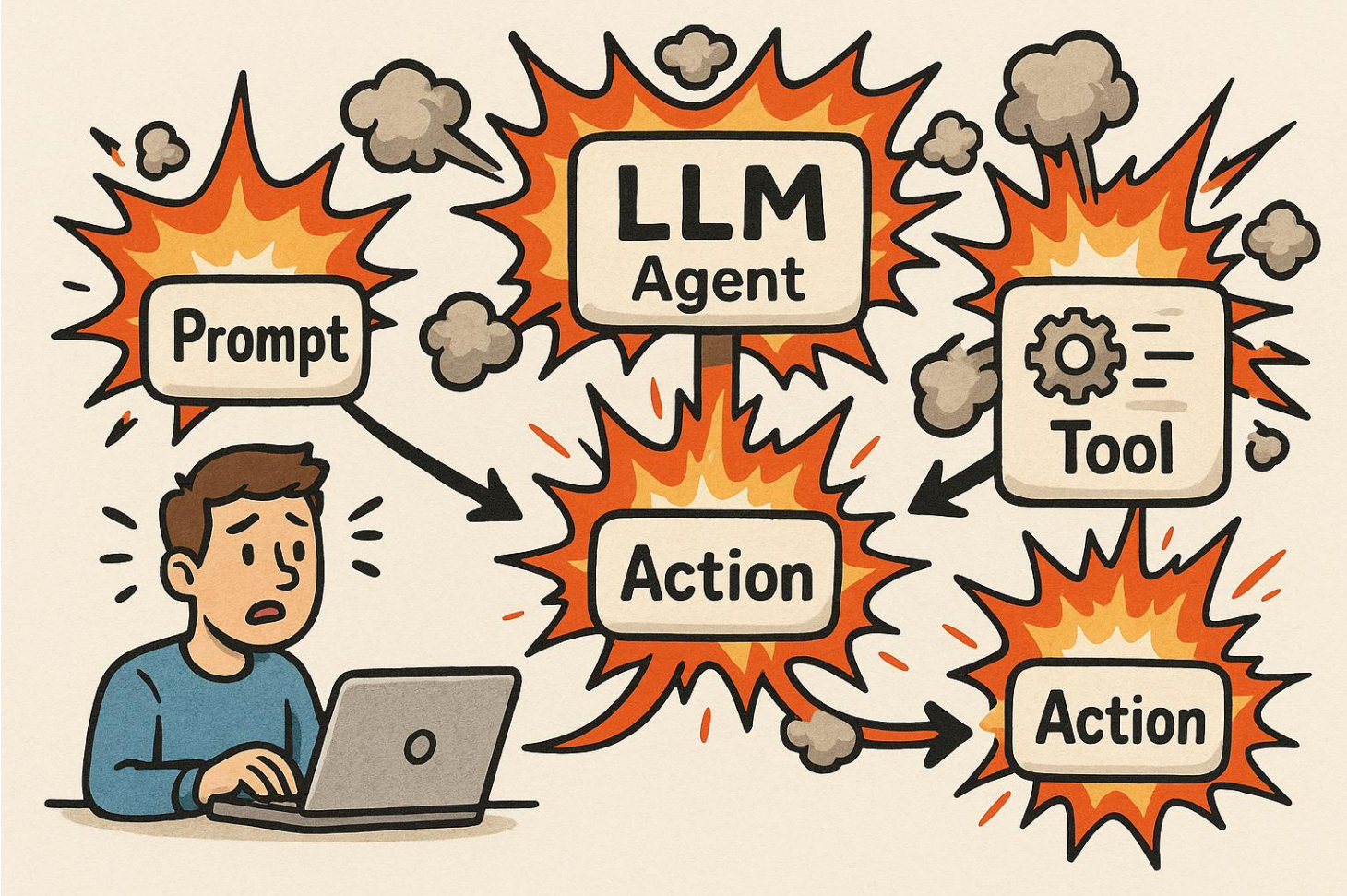
Please proofread the refined text for grammar, spelling, punctuation, terminology, and overall fluency. Ensure the translation accurately reflects the original meaning and style. Provide only the final, polished translation.

然而，他表示，削减开支最“负责任的方式”是通过全部12项法案。鉴于许多共和党人不愿看到政府停摆和国防开支削减1%，通过法案的紧迫性可能会促使足够多的共和党人与民主党人合作，延长四项支出法案的截止日期，并达成一项为政府提供资金的协议，即使这样做会让约翰逊与一些强硬的右翼人士产生摩擦。

Omniscien™ TECHNOLOGIES

Many Challenges

Brittle Workflows



- Hard to engineer
- Hard to maintain
- High cost due to inefficiencies
- Hard-code vs dynamically plan?
- LLM vs. tool use?
- Security risks

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

ADVANCE: 3

The Rise of Chinese AI and Models

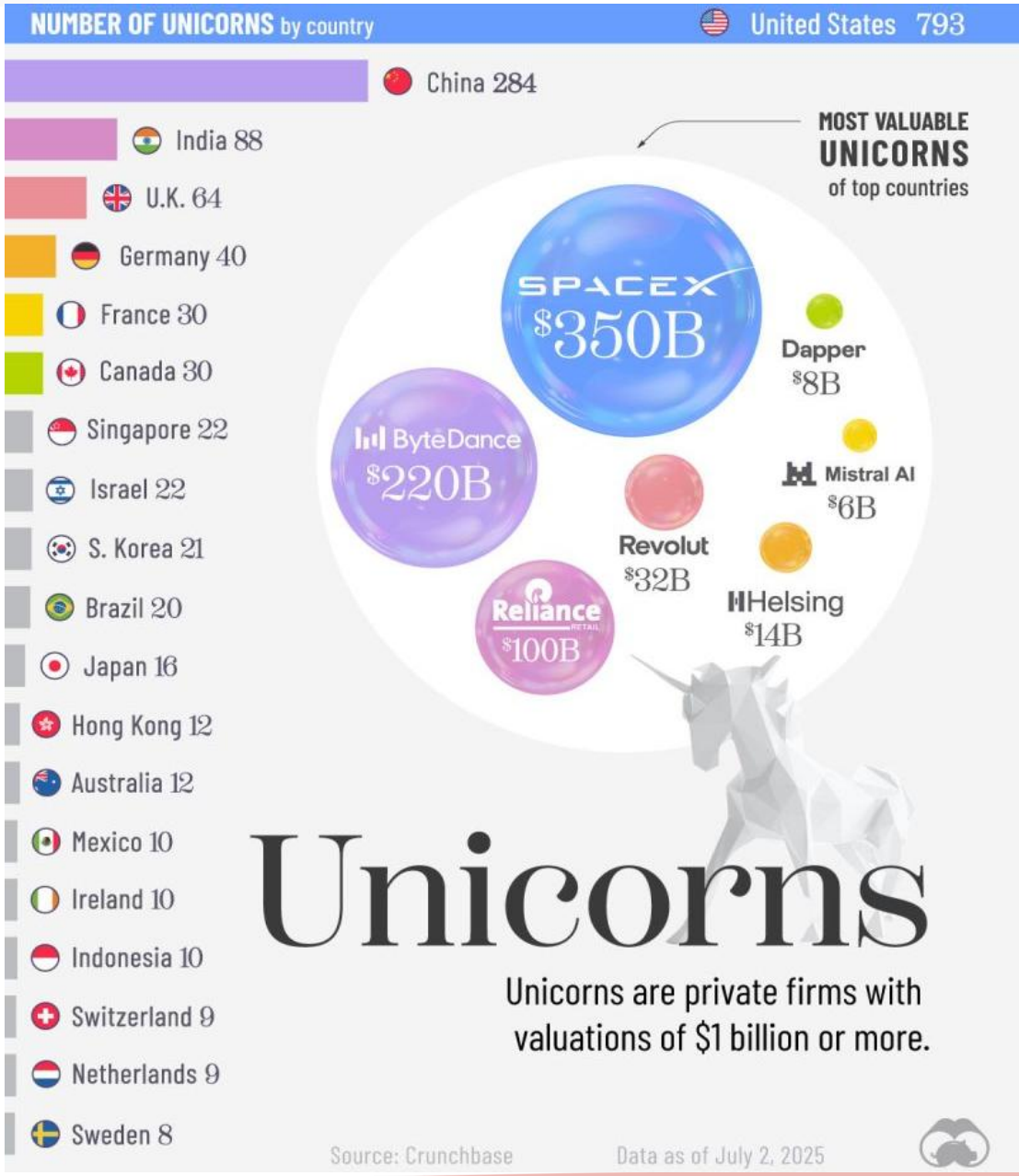
Cost, Scale, Efficiency

Global Unicorn Landscape: Power Is Concentrating, Not Spreading

The U.S. and China dominate the unicorn economy, holding over **1,000 combined**, while the next 20 countries together barely match a third of that.

AI and deep-tech unicorn valuations skew heavily toward scale, capital access, and data advantage, reinforcing the gap rather than closing it.

Smaller nations depend on fragmented ecosystems, making sovereign digital strategy essential if they want to avoid permanent dependency on U.S./China giants.



Omniscien™ TECHNOLOGIES

Why?

1. LLAMA is not truly open source, Chinese models are.

- Meta fails **9/10 of the OSI criteria**; LLAMA is “source-available,” not open. Chinese models (Qwen, DeepSeek, Baichuan, Yi) ship **real open weights, real redistribution rights, real commercial freedom**.

2. Cost-performance is unmatched: 30–70% cheaper compute for similar or better capability.

- Chinese model architectures are optimised for **training efficiency**, **context scaling**, and **lower inference cost**, giving startups runway they can’t get with US models.

3. True sovereignty and no vendor lock-in.

- Chinese open-weights can be **self-hosted, modified, retrained, fine-tuned, forked**, and deployed on any hardware — no API lock-in, no licensing traps, no kill-switch risk.

4. Accelerator independence — built for non-Nvidia hardware.

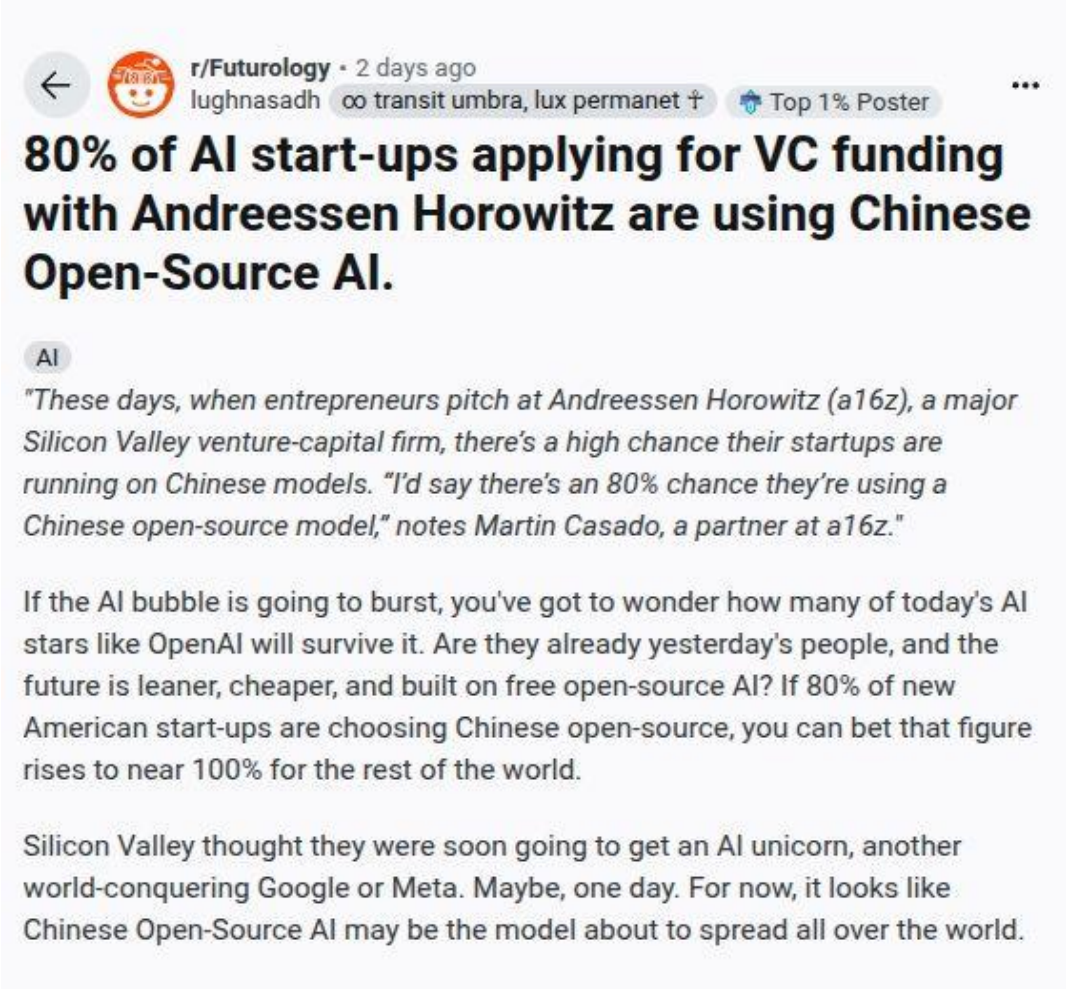
- Qwen/DeepSeek families run efficiently on **Huawei Ascend, domestic GPUs, FPGA accelerators, even consumer cards**, giving startups resilience against US GPU export controls.

5. Faster update cycles and brutally pragmatic engineering.

- Chinese labs iterate at **2–4× the velocity** of Western labs: faster releases, larger training runs, more variants, and more innovation at the edge and on-device.

6. Global-language strength: English parity + Asian, Middle Eastern, African languages where US models fail.

- Chinese models dominate **multilingual**, **low-resource**, and **non-Western cultural grounding**, making them far more relevant for global products.



r/Futurology • 2 days ago
lughnasadh ∞ transit umbra, lux permanet ☥ Top 1% Poster

80% of AI start-ups applying for VC funding with Andreessen Horowitz are using Chinese Open-Source AI.

AI

"These days, when entrepreneurs pitch at Andreessen Horowitz (a16z), a major Silicon Valley venture-capital firm, there's a high chance their startups are running on Chinese models. "I'd say there's an 80% chance they're using a Chinese open-source model," notes Martin Casado, a partner at a16z."

If the AI bubble is going to burst, you've got to wonder how many of today's AI stars like OpenAI will survive it. Are they already yesterday's people, and the future is leaner, cheaper, and built on free open-source AI? If 80% of new American start-ups are choosing Chinese open-source, you can bet that figure rises to near 100% for the rest of the world.

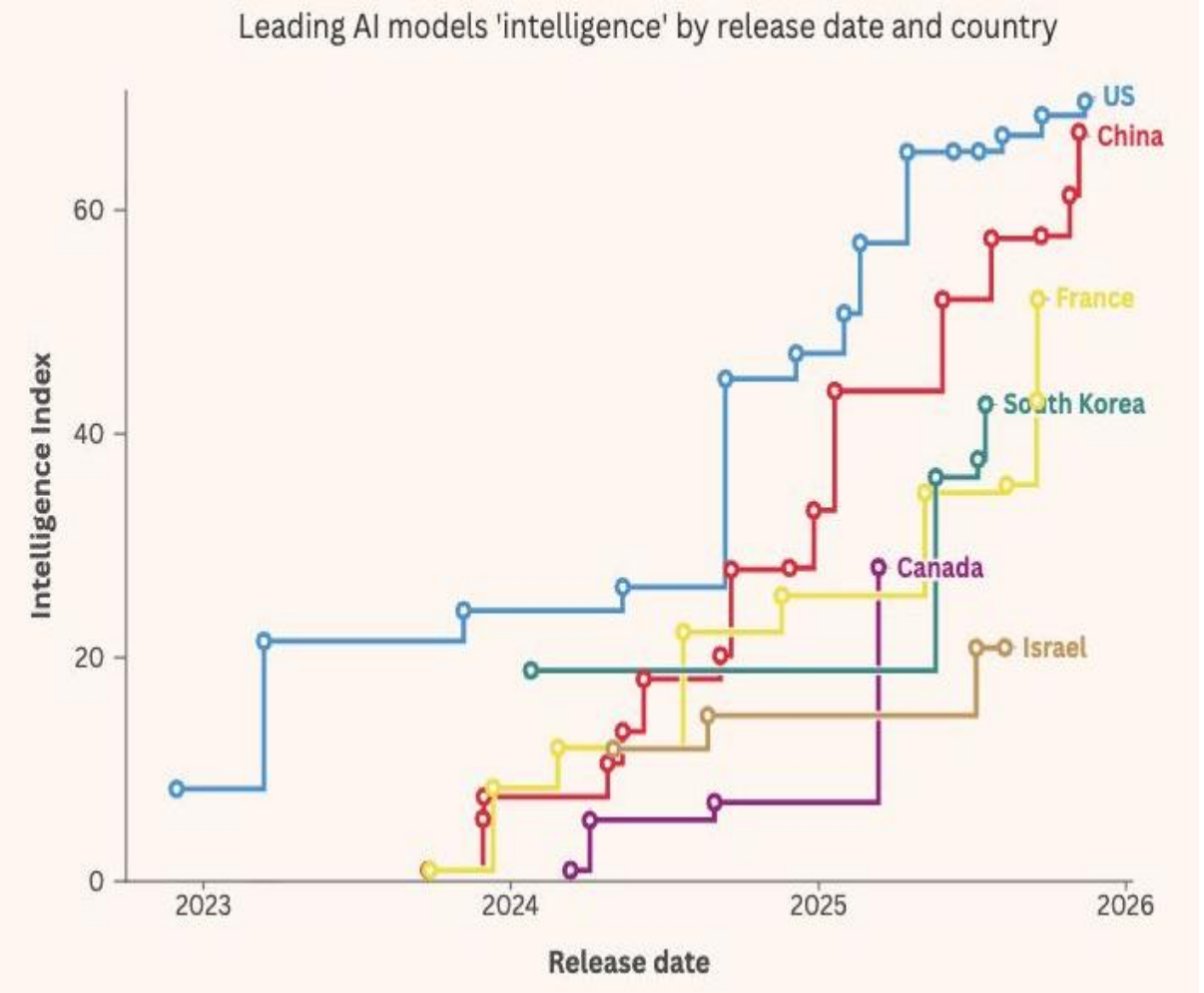
Silicon Valley thought they were soon going to get an AI unicorn, another world-conquering Google or Meta. Maybe, one day. For now, it looks like Chinese Open-Source AI may be the model about to spread all over the world.

Frontier AI Capability Gap Is Closing Faster Than Expected

China has almost erased the previous capability gap, matching or surpassing U.S. models across major benchmarks. The gap is now measured in months instead of years.

France, South Korea, Canada and Israel show steady improvement, but their progress remains gradual compared with the rapid jumps from the U.S. and China.

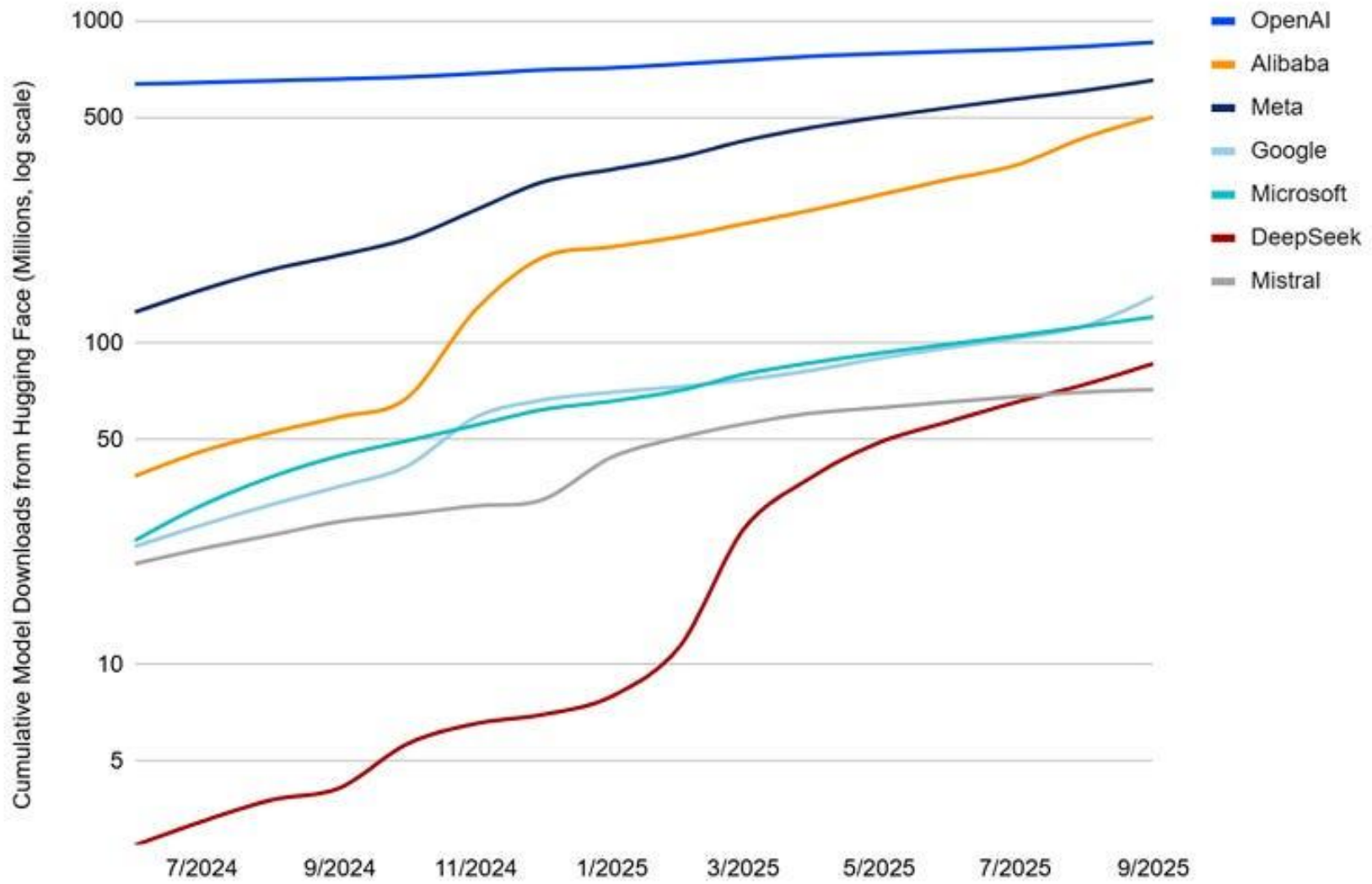
The race is now about release cadence, with fast iteration speed becoming the most important driver of frontier AI leadership.



Source: Artificial Analysis Intelligence Index

Omniscien™ TECHNOLOGIES

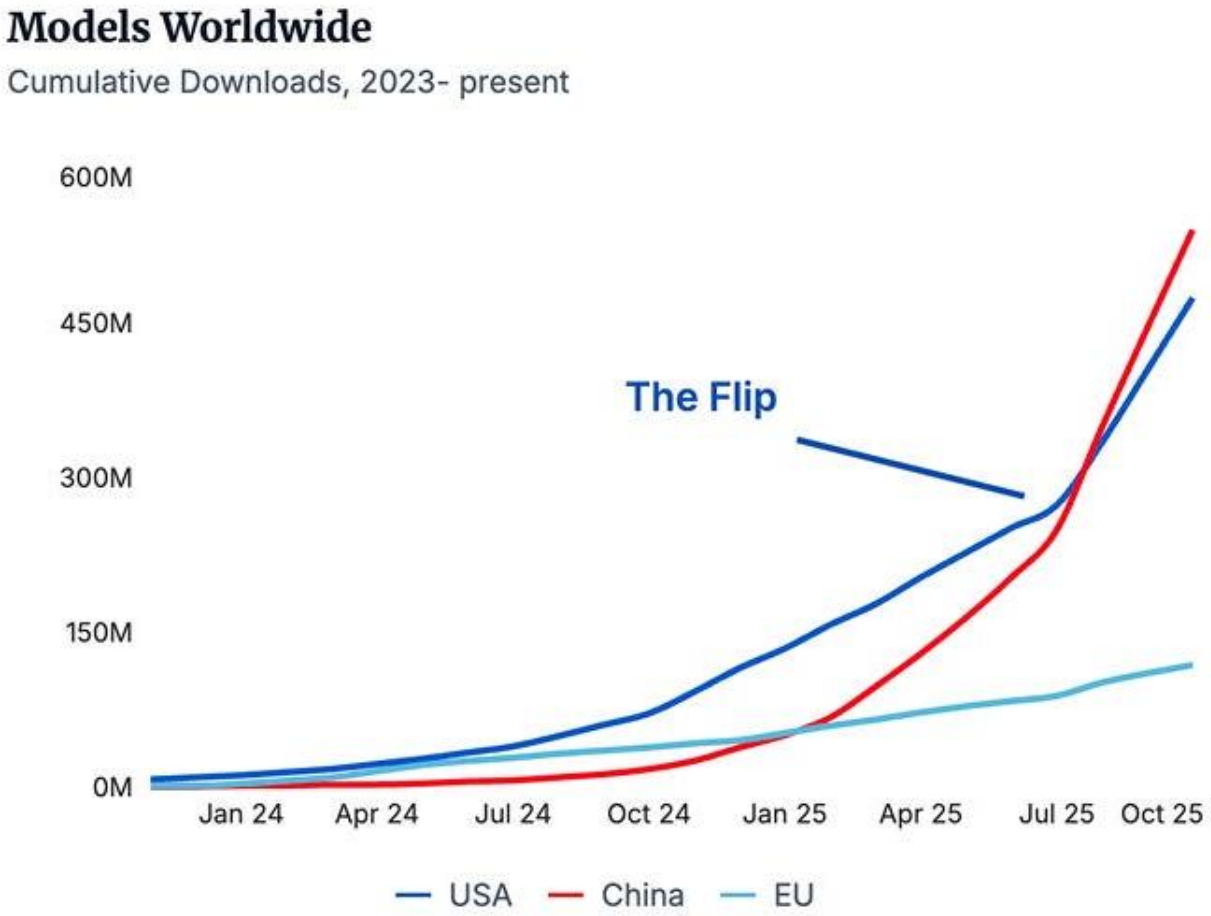
More Chinese Models than Western Models



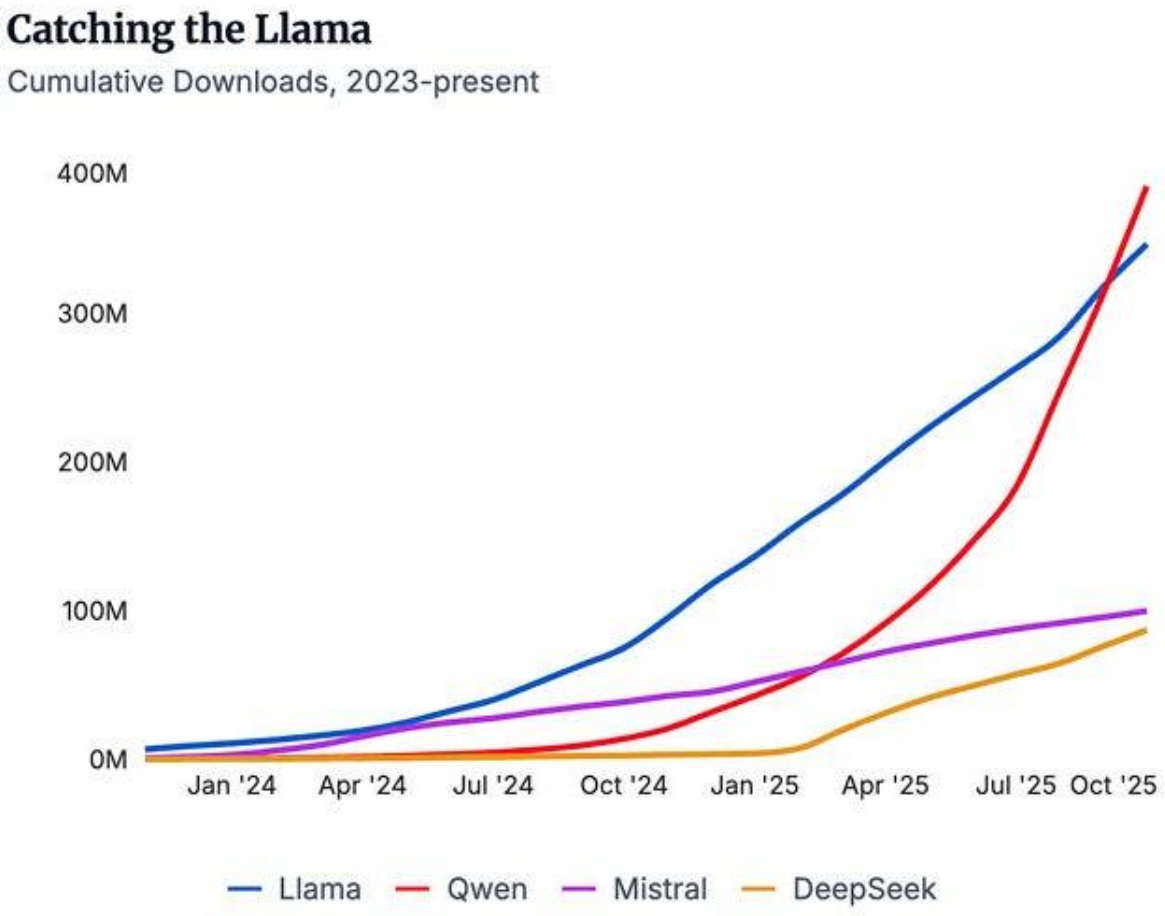
Source: Huggingface

Omniscien™ TECHNOLOGIES

More Chinese Models than Western Models



Source: Huggingface



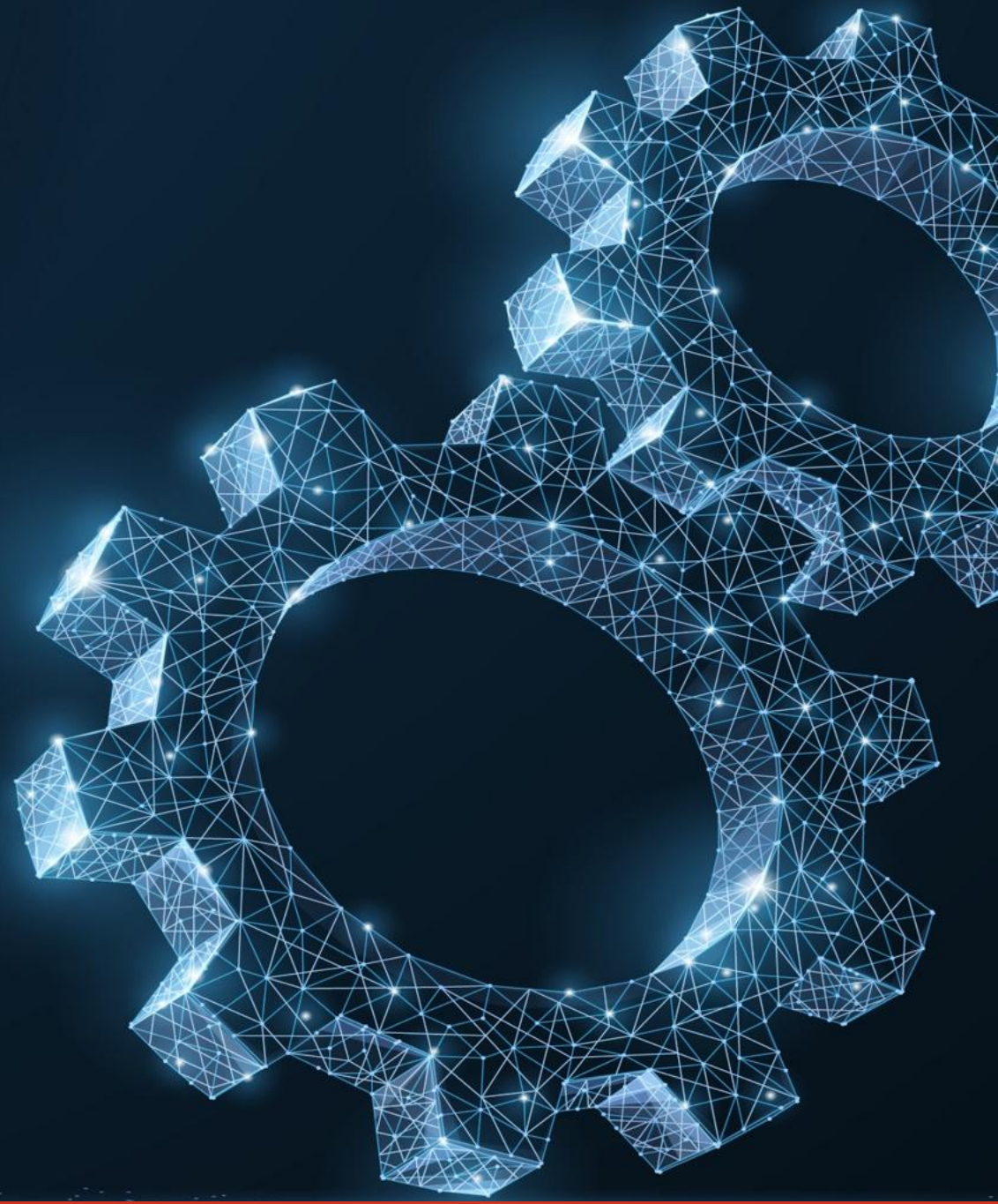
Source: Huggingface

Omniscien™ TECHNOLOGIES

ADVANCE: 4

On-Device AI

Private, Offline, Zero-Cost Inference



The Collapse Of Cloud Dominance

The biggest shift in AI is not model size. It is location. Compute is leaving the data center.

- Full LLMs now run on phones, laptops and edge devices.
- Inference cost falls toward zero when run locally.
- Enterprises bypass hyperscalers for operational AI.
- National infrastructure becomes resilient because workloads stay inside borders.
- Vendor lock-in breaks as every device becomes a sovereign inference node.
- Cloud outages no longer halt national capability when core functions run locally.

This is the end of centralised AI.

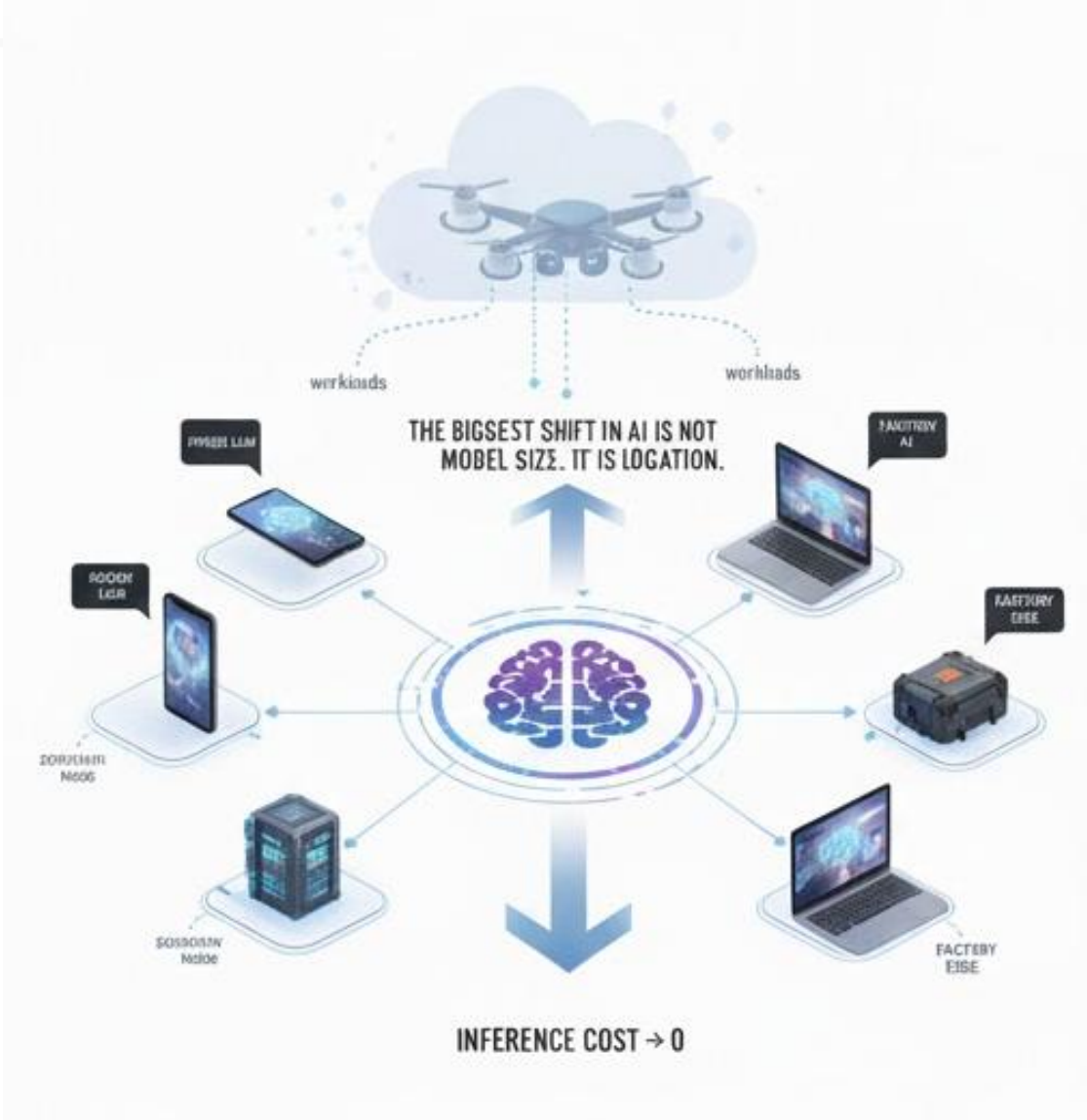
Omniscien™ TECHNOLOGIES

The Rise Of Sovereign Edge Compute

Power, data and compute are converging at the edge.

- China accelerates the shift because grid constraints force local compute.
- Industrial zones deploy micro data centers next to power generation.
- Edge clusters run domestic GPUs, ASICs and open designs without Nvidia bottlenecks.
- Workloads move to factories, ports, mines and city systems.
- Data no longer travels. Power no longer travels. The cloud becomes optional.
- Outages of Cloudflare, AWS or Azure no longer cascade across a nation.

Edge compute is sovereign compute.



Omniscien™ TECHNOLOGIES

The Personal AI Stack Revolution

AI stops being a cloud service. It becomes a personal asset.

- Every device becomes a private inference engine.
- Personal AI stacks run offline, private and unmonitored.
- Latency drops from hundreds of milliseconds to sub-10ms.
- Citizens gain digital autonomy because models never leave the device.
- Regulators gain safety because sensitive workloads stay local.
- Militaries, governments and critical operators run air-gapped AI by default.

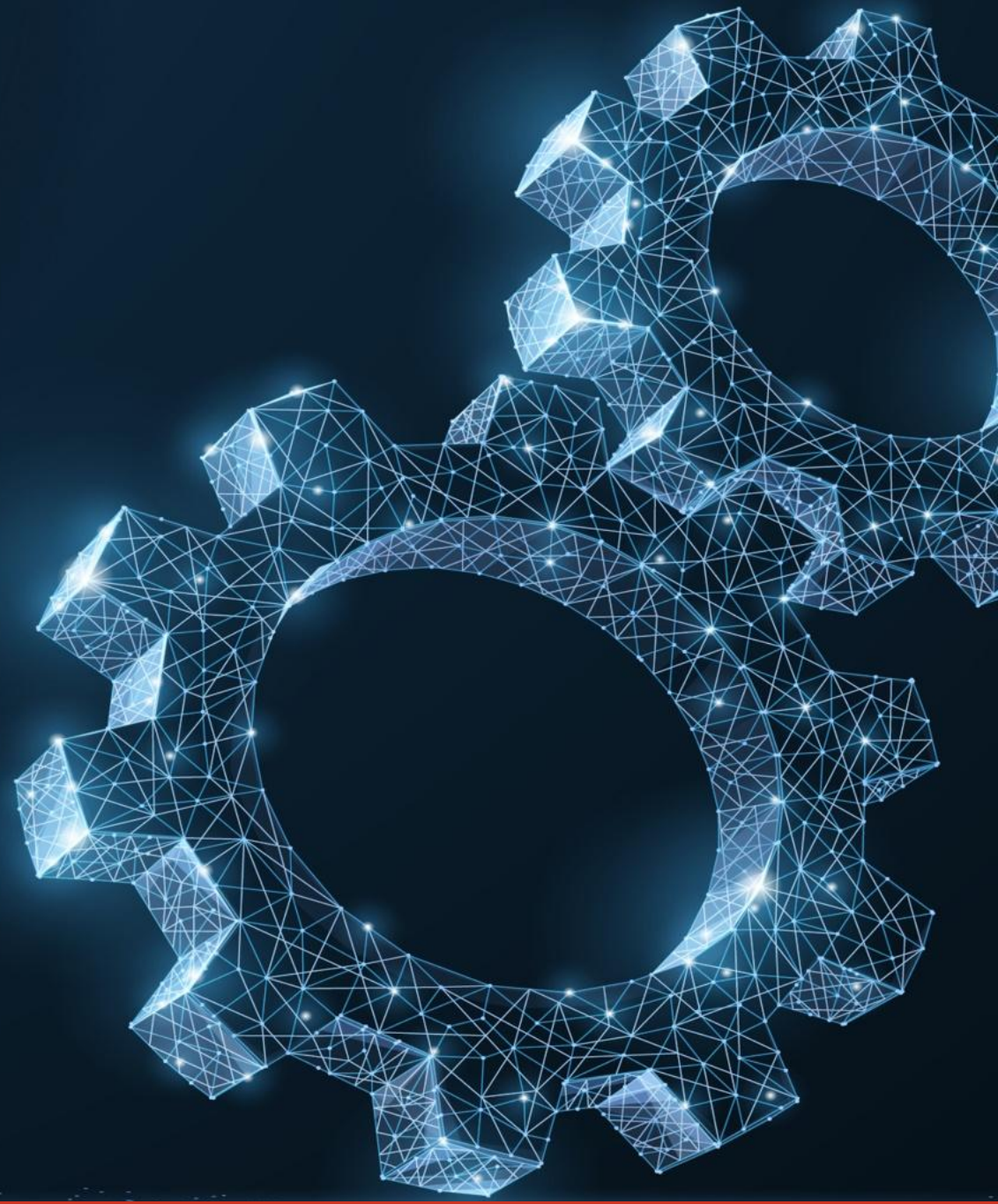
This is the most sovereignty enhancing shift of the decade.

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

ADVANCE: 5

Geopatriation
The Great Cloud Reversal



The Great Cloud Reversal Has Already Started

- New Gartner Term: **Geopatriation**

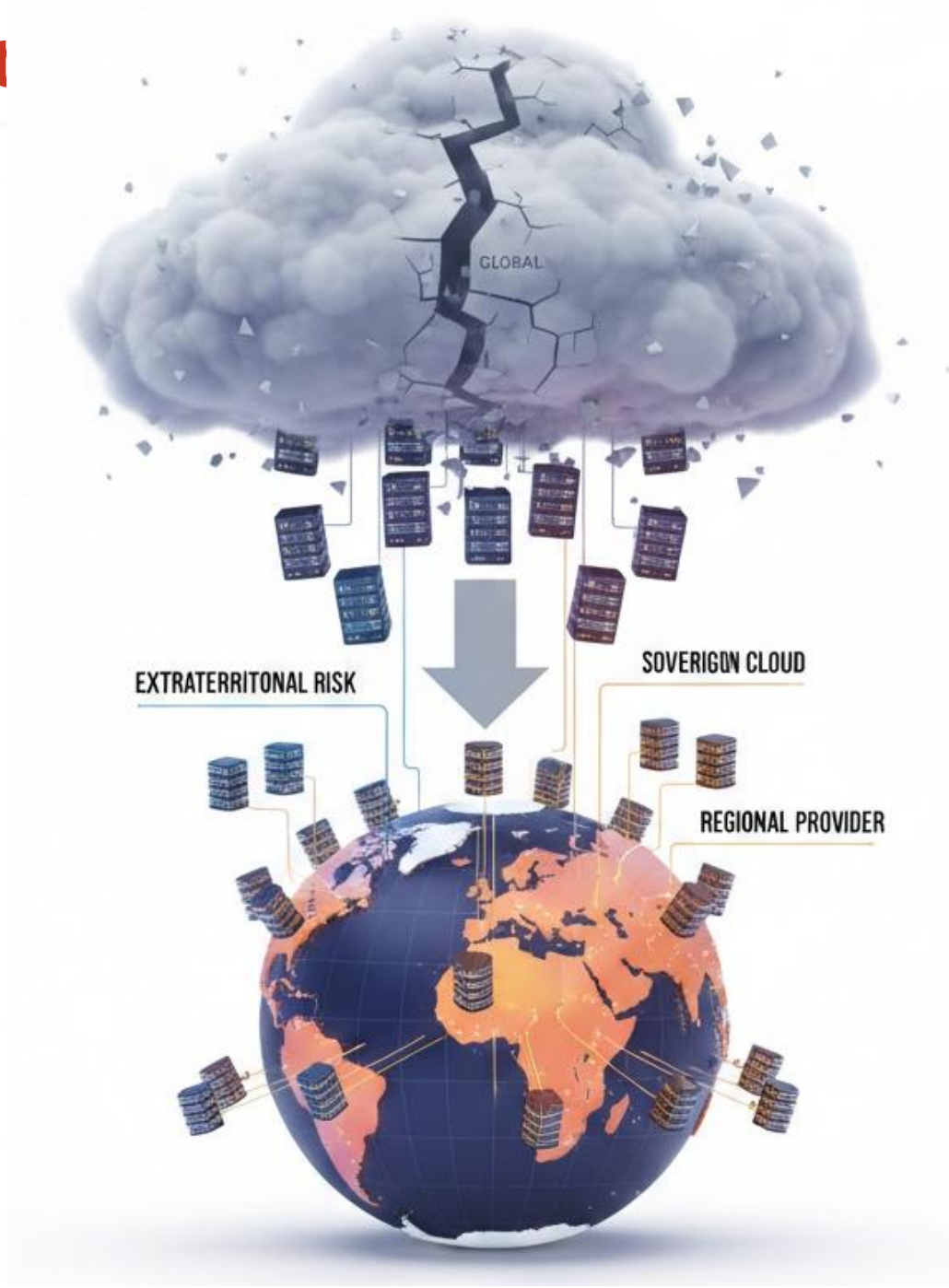
Geopatriation means moving company data and applications out of global public clouds and into local options such as sovereign clouds, regional cloud providers, or an organization's own data centers due to perceived geopolitical risk.

Cloud sovereignty, once limited to banks and governments, now affects a wide range of organizations as global instability increases.

Omniscien™ TECHNOLOGIES

The Great Cloud Reversal Has Already Start

- The key shift in enterprise architecture is geopolitical, not technical.
- Workloads leave global clouds as trust collapses under extraterritorial risk.
- Sovereign clouds and regional providers surge as control assets return on premise.
- This marks the rise of jurisdictionally aligned compute and the end of cloud by default.



Omniscien™ TECHNOLOGIES

Geopatriation Is A Survival Response

- Digital infrastructure is now recognised as national infrastructure.
- Boards classify cloud placement as a geopolitical exposure.
- AI workloads move first because models and data are strategic assets.
- More than 75 percent of EMEA workloads will localise by 2030.

Omniscien™ TECHNOLOGIES

The Global Cloud Is Splitting Along Geopolitical Fault Lines

- The borderless cloud collapses as nations reassert jurisdiction.
- Organisations pull compute back inside domestic legal perimeters.
- Sovereign clouds and controlled colocation become the backbone of resilience.
- The future is not multi cloud. It is multi jurisdiction.

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

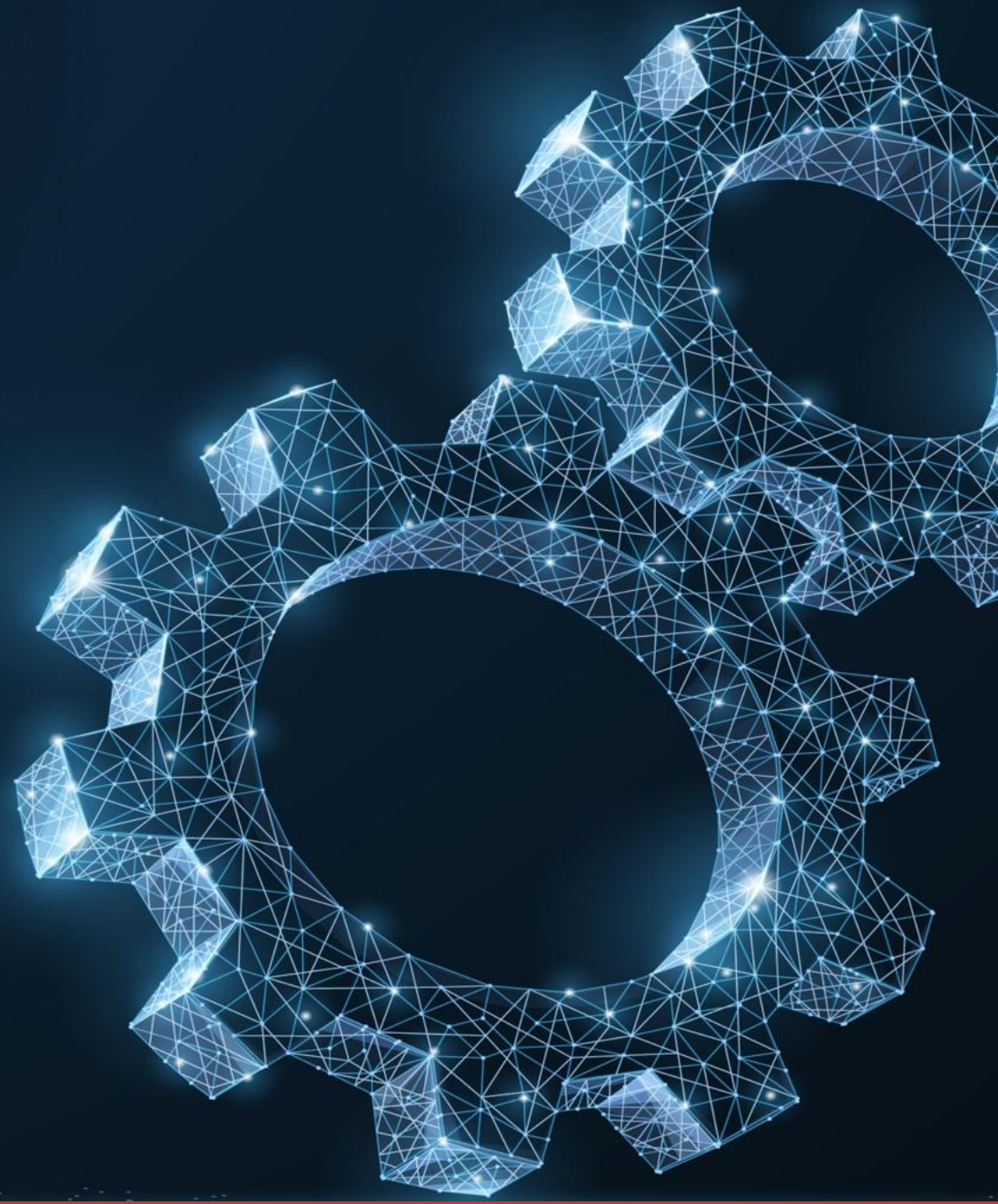
Top 5 AI Falsehoods that Most People Believe In

Omniscien™ TECHNOLOGIES

FALSEHOOD #1

95% of All AI Projects Fail to deliver ROI within 6 months

Is that a bad thing or a good thing?



MIT Study: 95% of GenAI projects fail to return ROI in 6 months

$$\text{Payback Period (Years)} = \frac{\text{Total Investment}}{\text{Annual Net Cash Flow}}$$

The media is in a frenzy

- No ROI in 6 months
- No tangible benefits
- POC Purgatory – Never Ending
- Shadow AI use outpaces official deployments
- Investment bias: chasing visible wins, missing real ROI.

Causes of failure include:

- Lack of **workflow integration** and **customization**
- **Brittle and inflexible** – constantly breaking
- Systems that **don't learn or retain context**
- Executive and senior management **discarding 50+ years of best practices** to **chase hype** and **cost savings** instead of **tangible business value**.

Familiar Pattern: If you replace 'GenAI' with 'blockchain' or 'big data,' the stats would be the same. No surprise.

Omniscien™ TECHNOLOGIES

Global ROI Norm

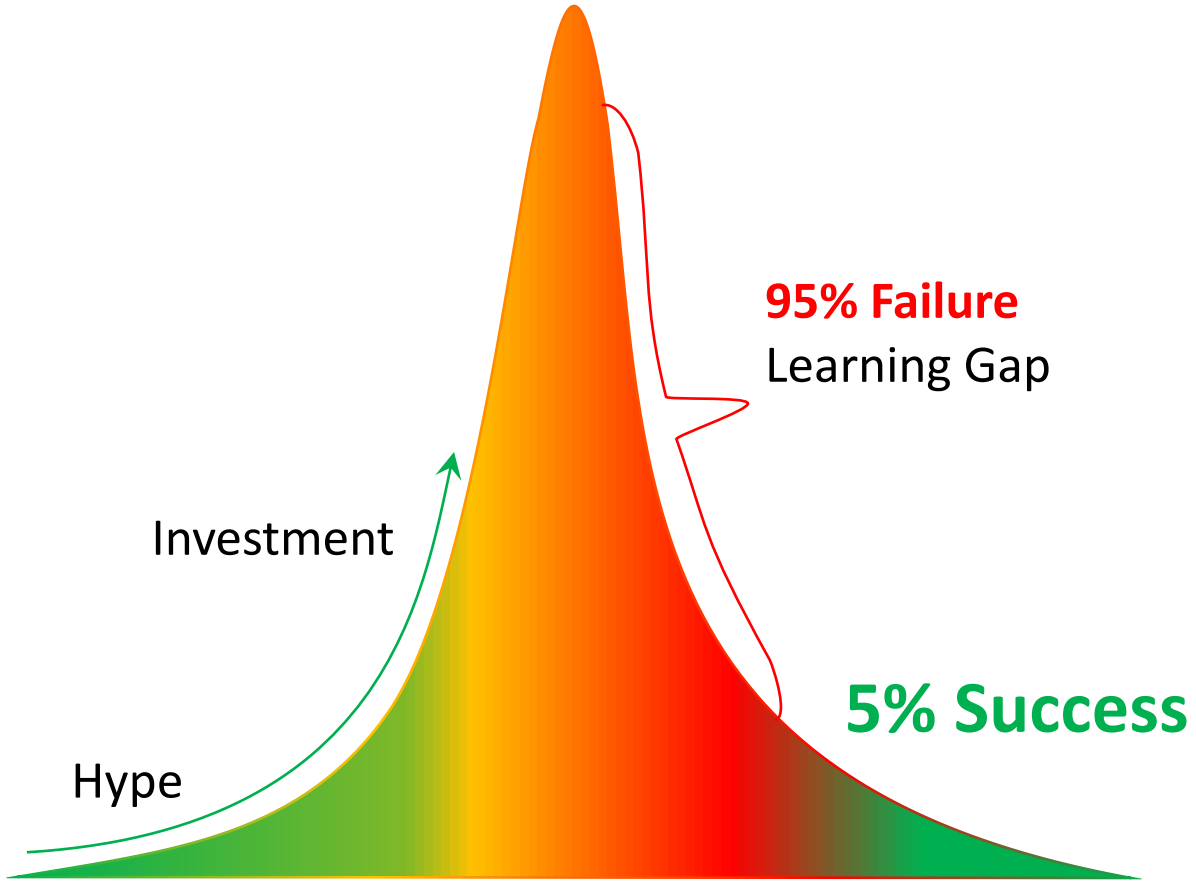
Standard Business Case (3 Years): A **three-year horizon** is the most common timeframe used for building and evaluating the full business case and ROI for major technology, capital, and organizational projects, allowing time to pass the initial implementation phase and reach steady-state returns.

Payback Period Target (1 Year or Less): Projects focused on operational efficiencies and cost savings are often expected to demonstrate a rapid **payback period** (time to break even) ideally within **one year** to be prioritized and considered high-value.

Long-Term Investments (5+ Years): Capital-intensive projects, such as infrastructure, manufacturing, or large-scale construction, typically require a **five-year or longer** projection period to fully realize and measure their intended ROI.

Omniscien™ TECHNOLOGIES

This presentation will demonstrate how any organization can be in that 5%



- Rapidly overcome the learning gap with adaptive, persistent solutions.
- Embrace People-First AI principles: design for real users, not demos. **Amplify human capital, don't replace it.**
- Multiple smaller improvements are better and lower risk than betting on massive transformational change.
- Prioritize the use case and measurable business value, not technology hype.
- Build **flexible**, **interoperable workflows** that integrate with current processes and adapt as needs change.
- Establish continuous feedback and improvement loops—systems must learn and adapt.
- Design from day one for **digital sovereignty** and **data privacy**—own your stack, control your data.

Omniscien™ TECHNOLOGIES

Putting Real Numbers to This

REALITY CHECK:

- 60-80% of **ALL** IT projects fail
- 70-88% of Digital Transformation projects fail.

- ROI in 6 months is unheard of for most projects of any kind.

- Thus – the 5% that succeed are by definition a huge success.

Omniscien™
TECHNOLOGIES

Success Is Not an Accident. It's Engineered.

Most AI pilots fail. The media focuses on the 95% who chase hype and lose.
But the real story—the only story that matters
—is the 5% who make AI deliver real business value, at scale, for real people.

What Do the 5% Do Differently?

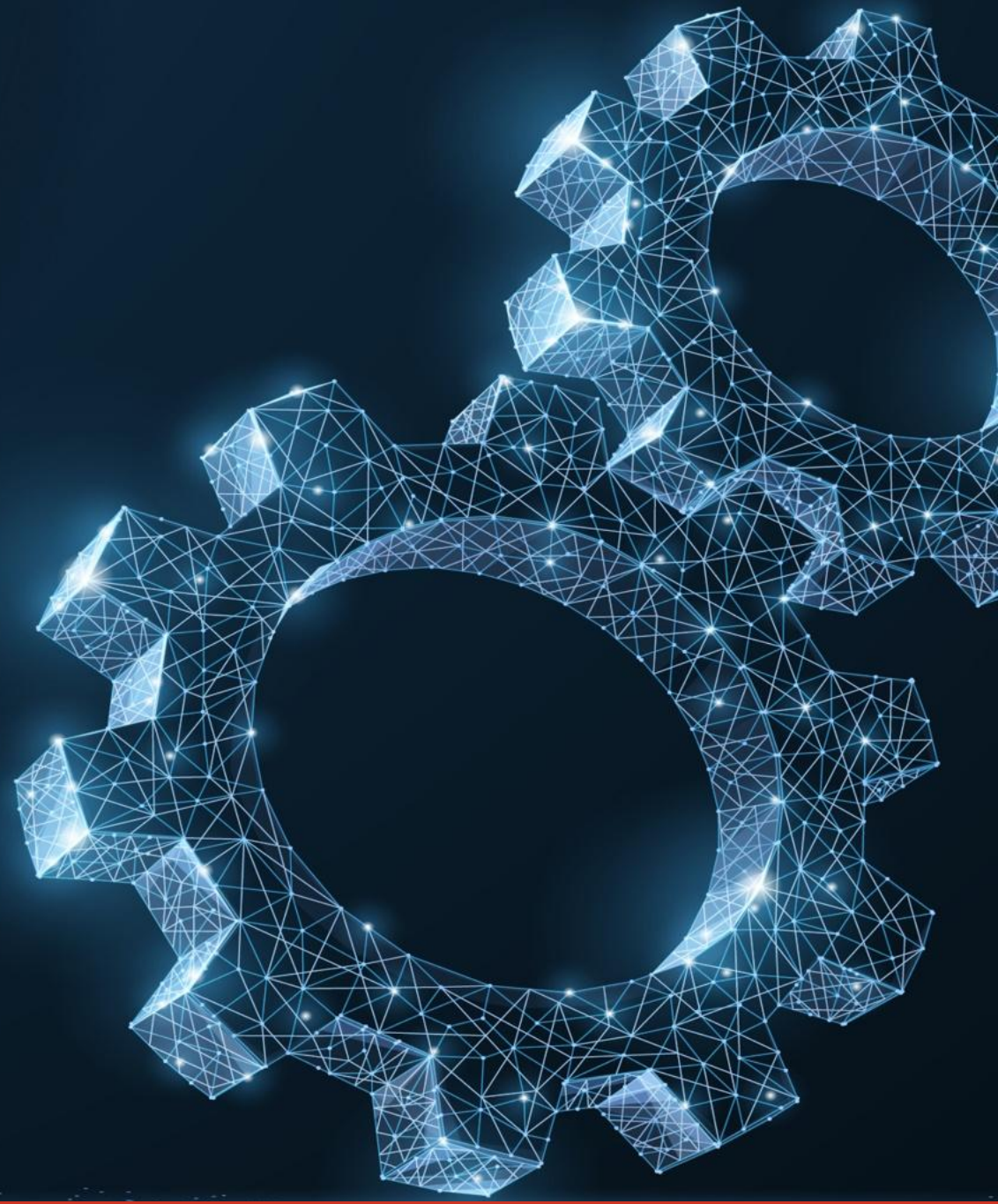
- Integrate AI directly into business workflows—where work actually happens, not as an afterthought
- Design for people, not just technology—amplify human skills, don't replace them
- Build for measurable ROI—focus on outcomes, prove value early, and keep iterating
- Own your stack and your data—make digital sovereignty a non-negotiable principle
- Establish continuous feedback and improvement—systems must learn, adapt, and get smarter over time

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

FALSEHOOD #2

Artificial General Intelligence (AGI) is here now, tomorrow, soon…



Predictions of AGI … Soon

- Extreme case: AI will take over the world and kill us all

- Reality
  - AI is a tool
  - Many fundamental flaws
  - We control how it is used

- Real Risks
  - Privacy, data security
  - Fake content, misinfo
  - Undermining education

AI 2027 [1]

Summary Research About

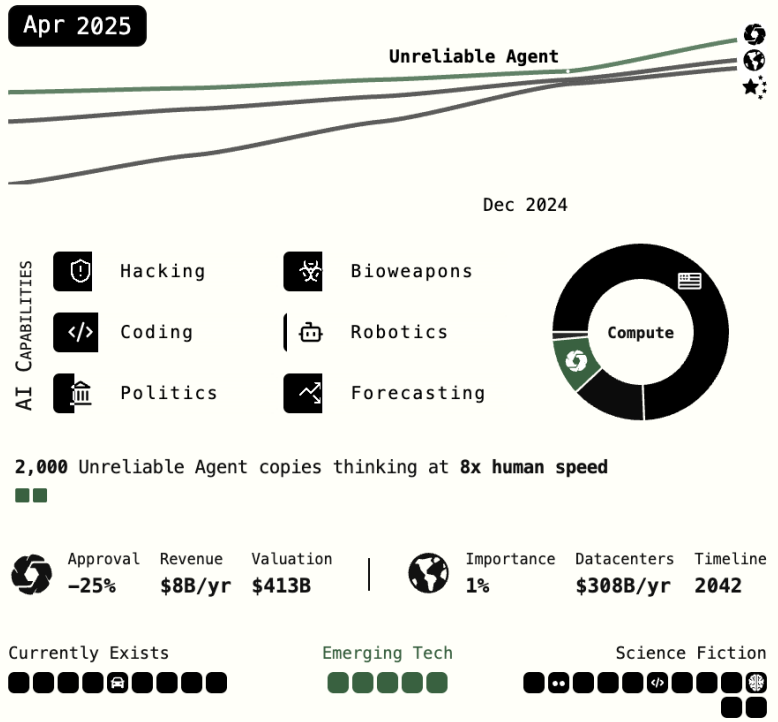
Daniel Kokotajlo, Scott Alexander, Thomas Larsen, Eli Lifland, Romeo Dean

We predict that the impact of superhuman AI over the next decade will be enormous, exceeding that of the Industrial Revolution.

We wrote a scenario that represents our best guess about what that might look like. It's informed by trend extrapolations, wargames, expert feedback, experience at OpenAI, and previous forecasting successes.[2]

What is this? How did we write it?

Why is it valuable? Who are we?

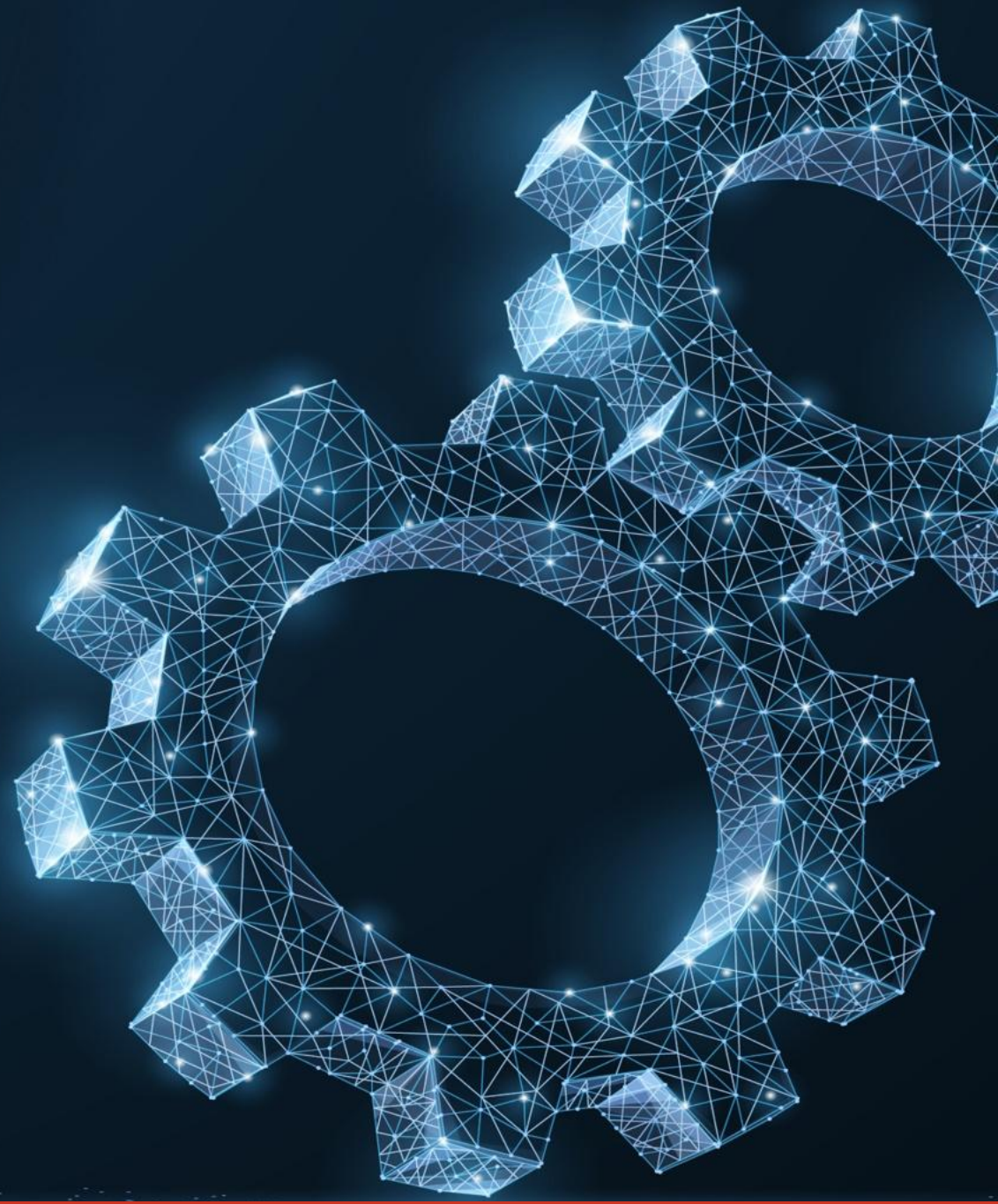


Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

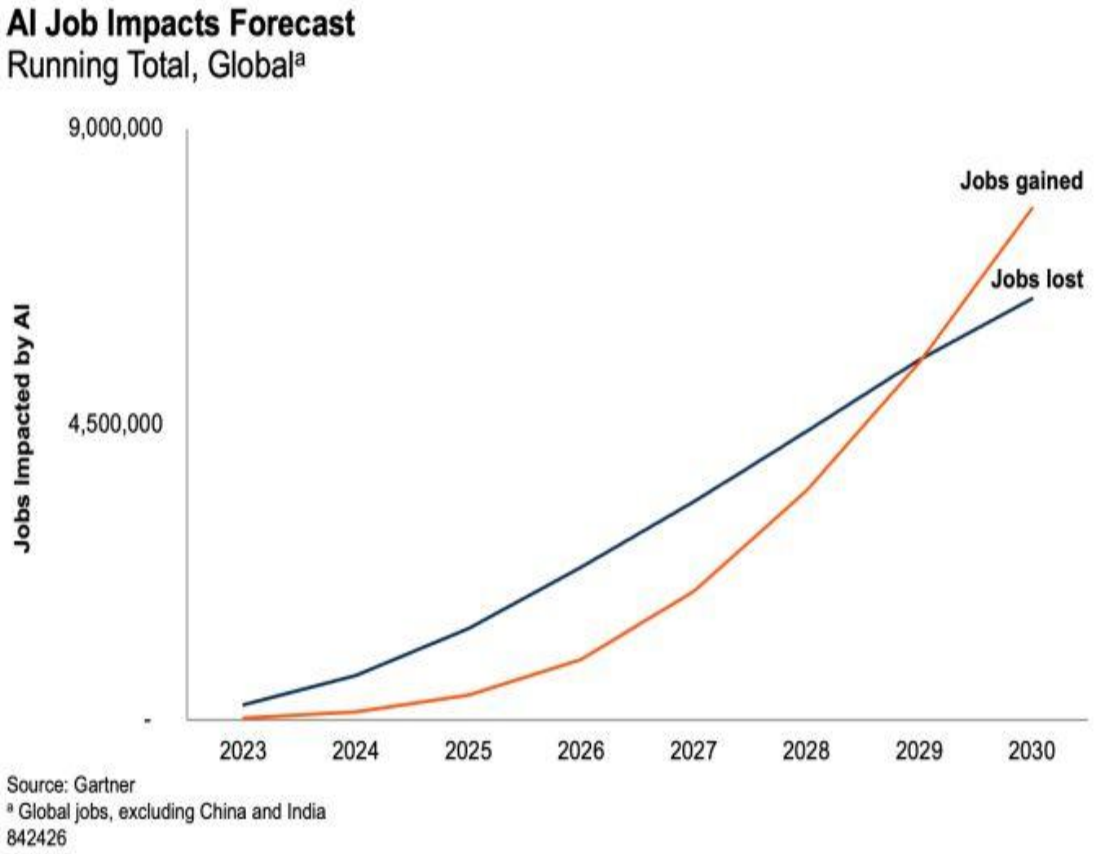
FALSEHOOD #3

AI Is Replacing Jobs at Scale



Reality: The Data Shows the Opposite

- Labour data across the US, EU and Asia shows no net job decline in AI-exposed sectors; employment rises where AI adoption is highest.
- Automation removes tasks, not roles; jobs restructure and workers shift to higher-value functions instead of disappearing.
- Firms that adopt AI expand, increasing demand for analysts, operators, reviewers, data specialists and integration roles.
- The real job losses come from companies that fail to adopt AI; organisational inertia destroys more roles than automation.
- AI boosts productivity and historically productivity increases correlate with hiring, not contraction.



Omniscien™ TECHNOLOGIES

AI Creates New Layers of Work, Not Unemployment

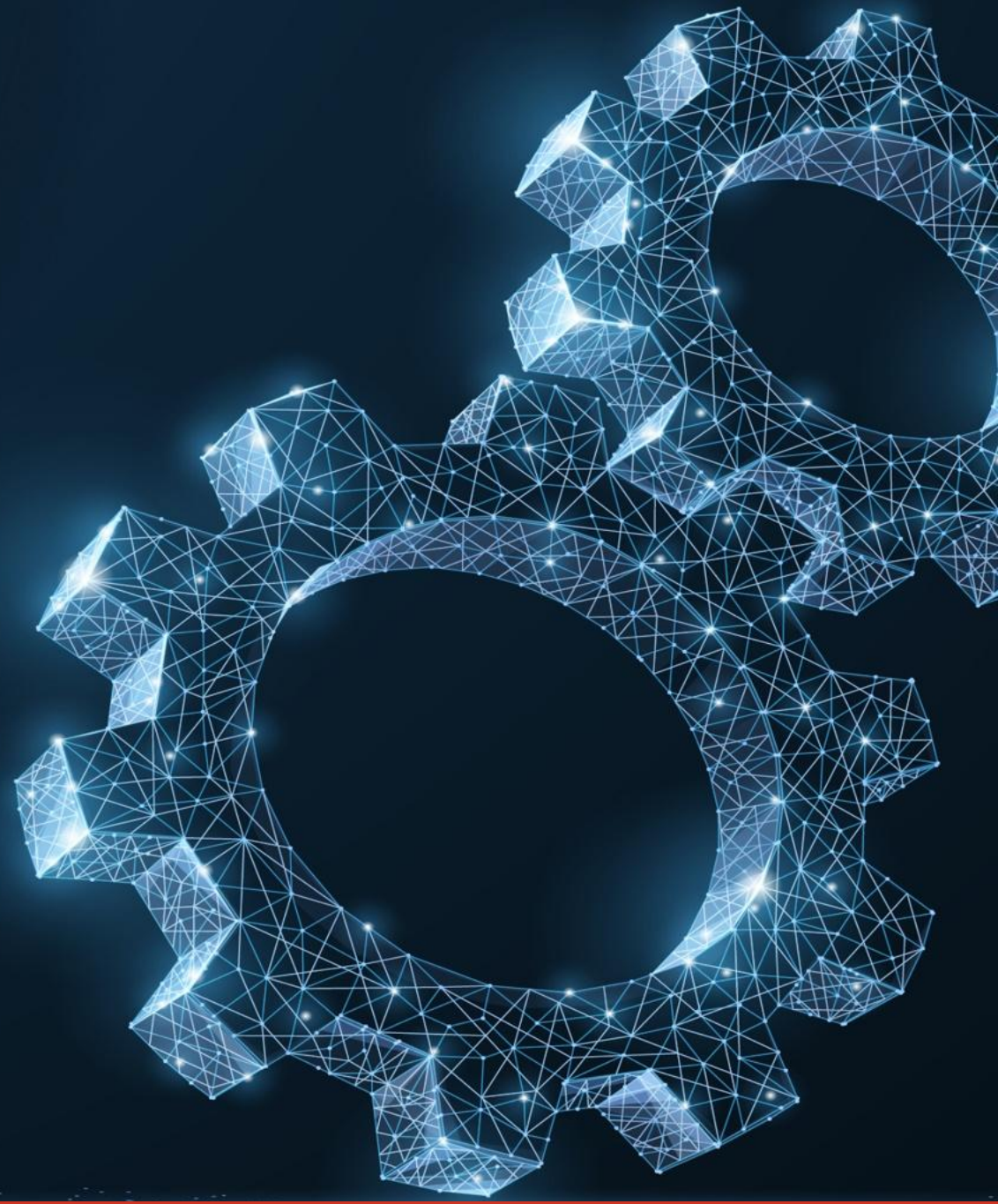
- New operational layers emerge: prompt engineering, AI auditing, workflow design, safety review and domain-specific agent management.
- Small firms hire fastest because AI reduces the cost of scaling and lets them compete directly with incumbents.
- On-device and edge AI create roles in cybersecurity, MLOps, embedded systems, localisation and private deployments.
- AI removes low-value copy-paste work, not people; staff get redeployed into product, engineering, customer and revenue teams.
- Productivity-augmented workers earn more, stay longer and become more valuable, contradicting the “AI replaces humans” narrative.

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

FALSEHOOD #4

AI Hallucinations Will Soon Be Fixed



Reality: They Cannot Be Eliminated

- LLMs predict the most probable next token, not the most true one; probability, not accuracy, drives the model.
- Hallucinations arise from incomplete, conflicting or low-signal training data; they are not a patchable bug.
- Even perfect data cannot stop hallucinations because prediction models generalise beyond inputs by design.
- Scaling reduces frequency but increases confidence, making hallucinations harder to detect and correct.
- Models invent structure when none exists because the objective is coherence, not factual verification.

Omniscien™ TECHNOLOGIES

Containment Is Possible. Elimination Is Not

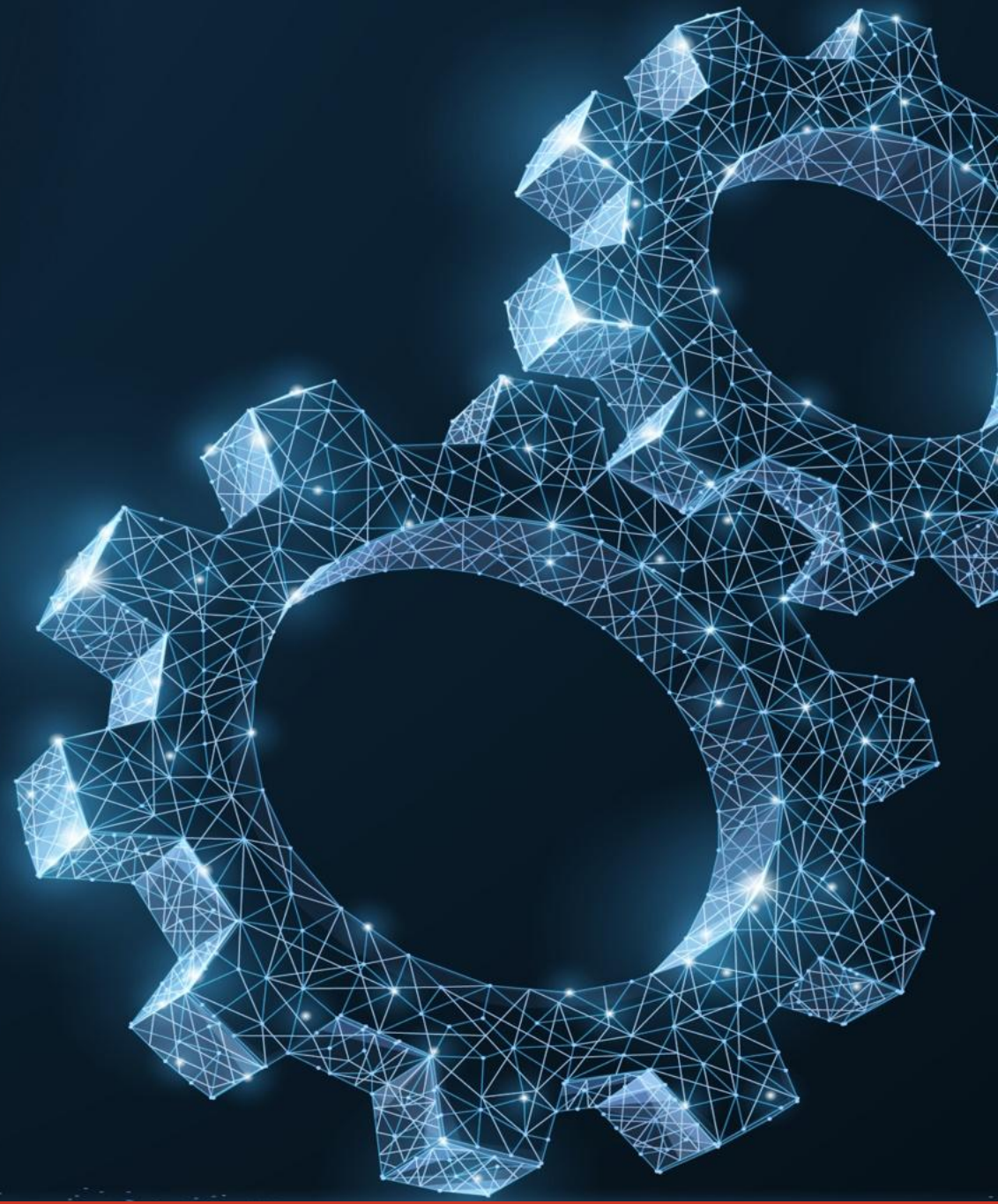
- Retrieval reduces hallucinations but also imports errors when sources are wrong, outdated or low quality.
- Agentic systems hallucinate more because multi-step reasoning compounds small errors into large false outputs.
- Deterministic systems do not hallucinate but lack LLM flexibility, making hybrid architectures the only viable approach.
- Verification layers work: audits, guardrails, cross-model checks, human review and reference validation.
- Any vendor claiming zero hallucinations is selling marketing, not science; the mathematics make that impossible.

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

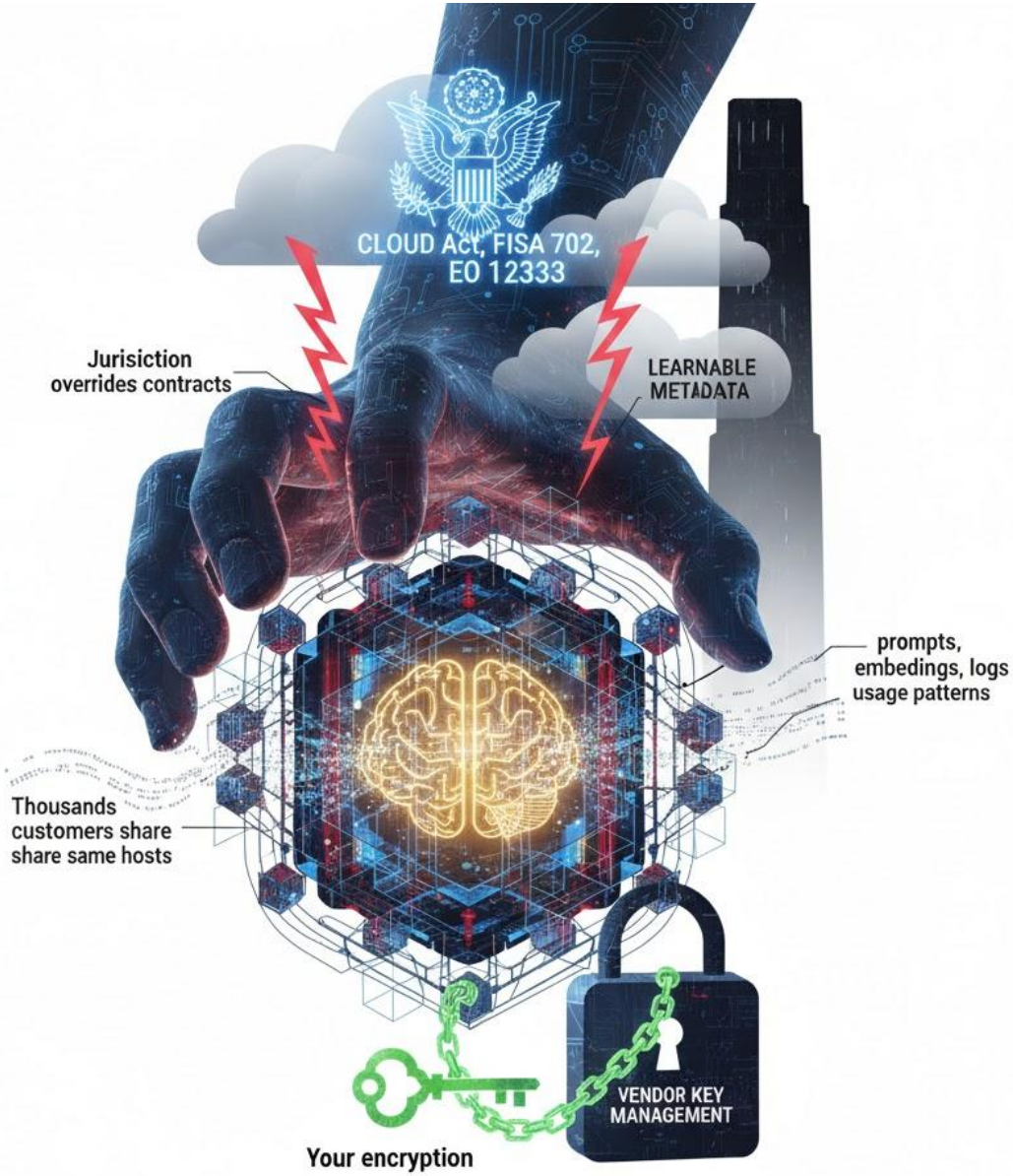
FALSEHOOD #5

Your Data Is Safe in OpenAI, Google, and Other LLM and Cloud Services.



Reality: Control Is Not Custody

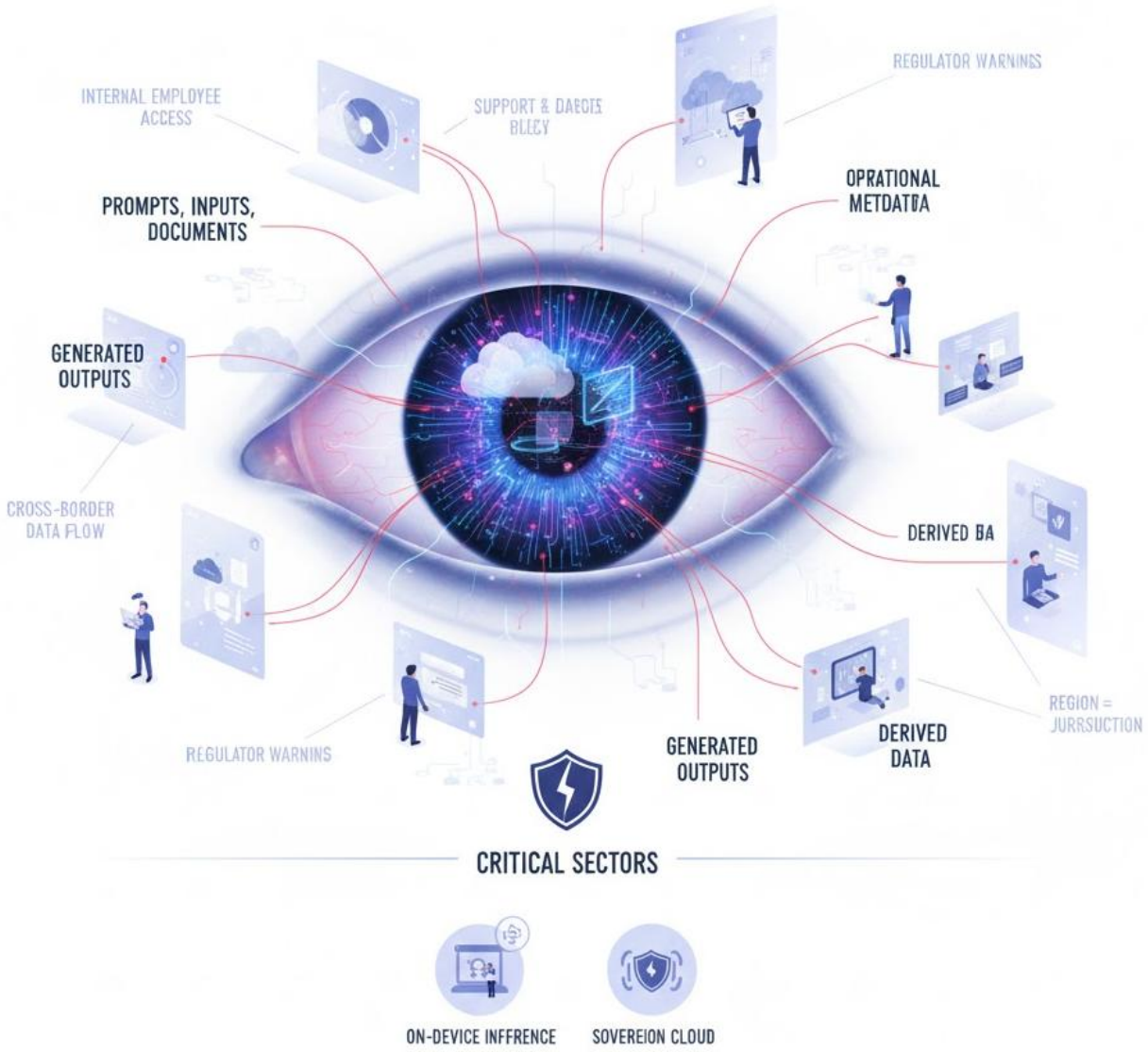
- Jurisdiction overrides contracts; CLOUD Act, FISA 702 and EO 12333 supersede your privacy terms the moment your data touches US-controlled infrastructure.
- Models still ingest operational metadata; prompts, embeddings, logs and usage patterns remain learnable even when content is excluded from training.
- Retention windows are opaque; telemetry, traces and system logs persist far longer than customers expect or are told.
- Multitenancy is structural; hyperscaler AI runs thousands of customers on the same hosts, GPUs, kernels and orchestration layers.
- Your encryption is not yours; key management and root-of-trust remain under vendor control, not customer control.



'Control Is Not Custody.

Omniscien™ TECHNOLOGIES

Cloud AI Sees More Than Customers Realise

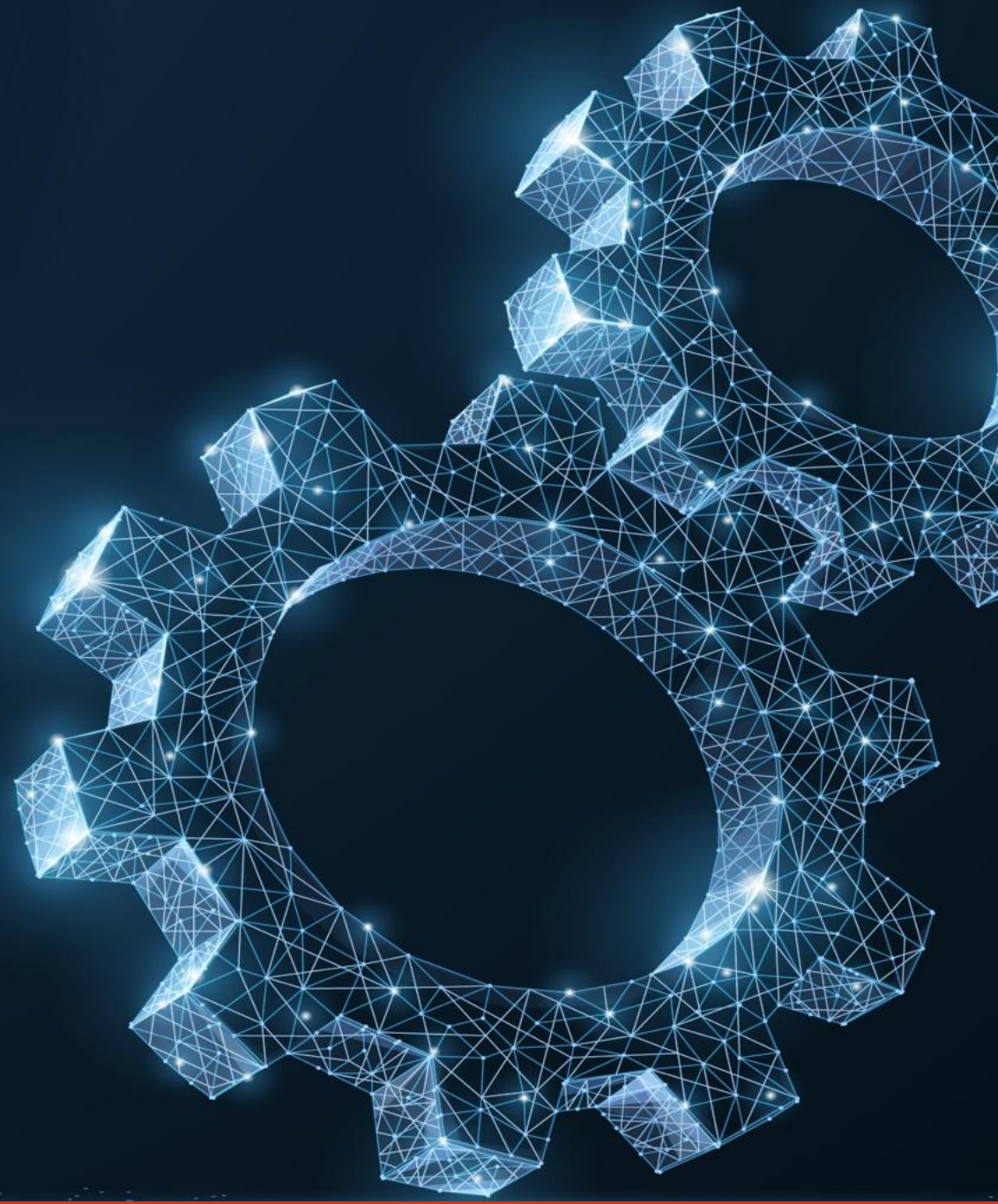


- LLM inference requires content visibility; providers see prompts, inputs, documents and generated outputs by design.
- Internal employee access is real; support, debugging and incident response pathways bypass customer-level controls.
- Data crosses borders inside provider backbones; region does not equal jurisdiction and cannot guarantee data locality.
- Model outputs create sensitive derived data that exposes patterns, relationships and business logic even when raw data is hidden.
- Regulators warn hyperscaler AI is not sovereign or private for critical sectors without isolation, on-device inference or sovereign-cloud alternatives.

Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

AI Snapshot for the Language Services Industry (LSPs)



The State of AI in Language Services

The Promise

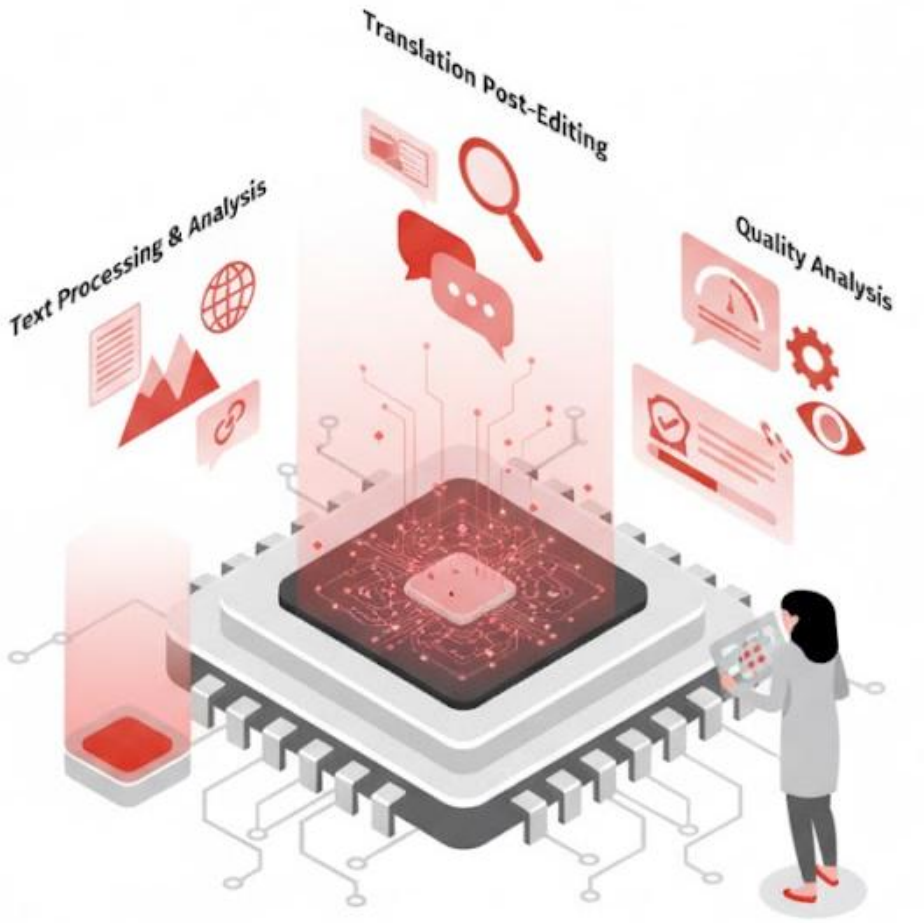
- Translation, subtitling, data processing at global scale
- Instant access, cost reduction, workflow acceleration

The Reality

- Hype often exceeds delivery
- Myths: “AI replaces humans” / “perfect translation is here”
- True strength: AI as a force multiplier, not a substitute

Industry Snapshot

- Widespread adoption across LSPs, enterprises, and media
- But uneven results: quality gaps, governance risks, sovereignty challenges



AI in language services is transformative — but only when fact is separated from fiction.

Omniscien™ TECHNOLOGIES

AI Capabilities: The Real Story

Text Processing & Analysis

- Handle billions of words at speed
- Multilingual ingestion & classification
- Pattern detection beyond human scale

Translation & Post-Editing

- High efficiency in bulk translation
- Consistent quality scoring & TAUS/TAQA benchmarks
- Blind spots: nuance, cultural trust, ethics

Quality Analysis

- Automated scoring accelerates workflows
- Still requires human oversight in sensitive domains

AI is not a replacement. It excels at scale, speed, and structure, but humans remain essential for nuance and trust.

Omniscien™ TECHNOLOGIES

Localization Revenue Reality Check

Reported industry surveys:

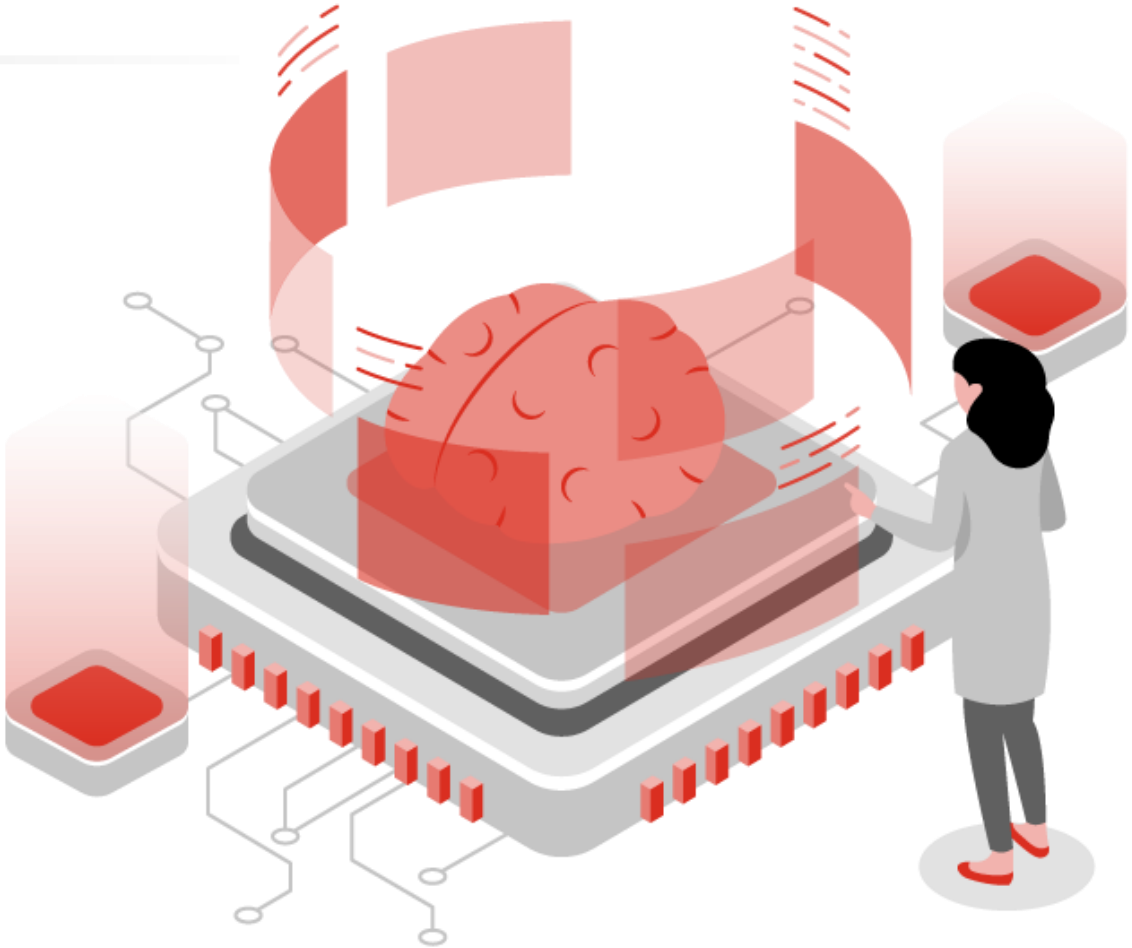
- “flat to slightly growing revenue”

Hidden truth:

- growth is **largely *data annotation/collection*** (~15–25%), not core localization

Result:

- fewer localization PMs and engineers needed year by year
- Revenue more stable in regulated/eLearning verticals, but deal sizes + active clients dropping in mid-tier LSPs



We are watching a structural decline in traditional localization revenue masked by other adjacent services.

Omniscien™ TECHNOLOGIES

The Economics of Tools vs. Revenues

MS pricing models have shifted:

- Usage-based fees
- Mandatory top-ups

Phrase and others:

- Per-word cost inflation
- Even as per-word LSP revenue falls

Then vs. Now:

- TMS costs rose from ~5% → >13% of linguist costs

Case study:

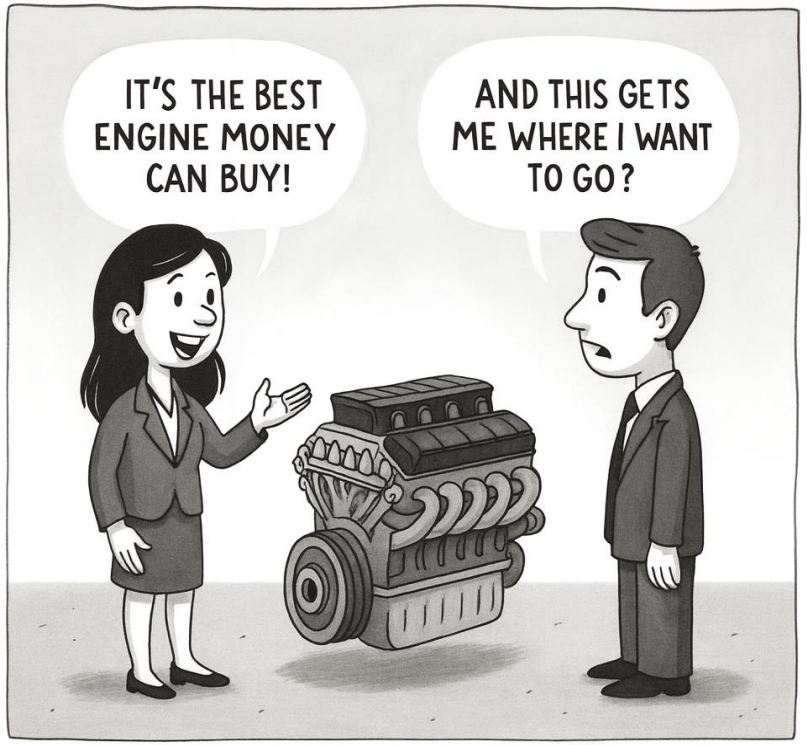
- Mid-sized LSP now pays as much annually for Phrase as a full ERP (Oracle Netsuite)

The technology layer is eating more margin than ever before.
The tools that were supposed to scale efficiency are now draining profitability.

Omniscien™ TECHNOLOGIES

LLMs are not a replace all, they are part of the solution

- **LLMs / Engines** – raw capability but incomplete on their own
- **Domain Models** – tuned for accuracy in specific sectors
- **Data Pipelines** – ingestion, cleansing, normalization, and governance
- **Human-in-the-Loop** – linguists, editors, and subject experts
- **Workflow Orchestration** – TMS, connectors, automation layers
- **Security & Sovereignty** – deployment control, data compliance, privacy
- **Evaluation & QA** – continuous testing, benchmarks, recovery proofs
- **Integration & APIs** – interoperability with existing systems
- **User Training & Change Management** – ensuring people know how to use AI responsibly



"Wheels, brakes, steering, and a small amount of common sense sold separately."

LLMs are powerful, but without data, governance, orchestration, and trained users, they're just the engine, not the car.

Omniscien™ TECHNOLOGIES

Large vs. Small Language Models

Trade-offs

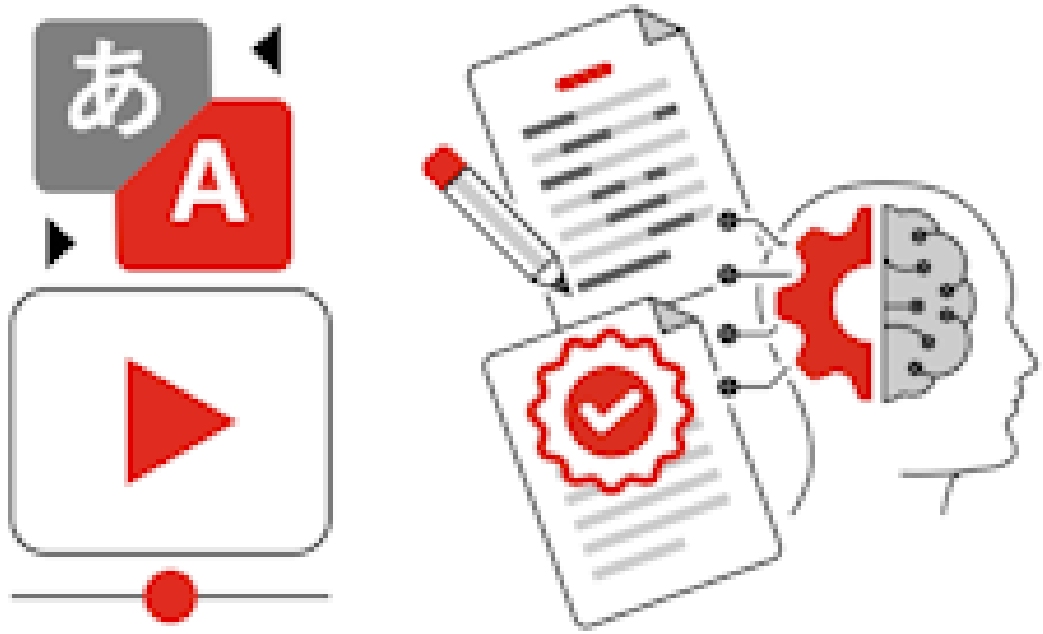
- Large LLMs → general knowledge, broad capabilities, high cost, less control
- Smaller domain models → faster, cheaper, higher accuracy in specialized tasks

Surprising Trends

- Domain-tuned small models often outperform giant LLMs in accuracy
- Lower compute & storage requirements = more sovereignty & control
- Cost/performance curve now favors fit-for-purpose over size for size's sake

Amplification

- AI as a force multiplier: scaling humans, not replacing them
- Combining models → best of both worlds (LLM reasoning + domain accuracy)



Bigger isn't always better. Sovereign, specialized models often win in practice.

Omniscien™ TECHNOLOGIES

Inclusivity & Low-Resource Languages

The Challenge

- Global language inequality: dominant languages get AI investment, others are ignored
- Low-resource languages = lack of training data, models, and tools
- AI models replicate existing market biases

Approaches

- Synthetic data generation & transfer learning
- Community-driven corpora and public initiatives
- Domain-specific small models for underserved languages

Case Studies

- Governments & NGOs using AI to preserve minority languages
- Enterprises adopting inclusive workflows to reach new markets

Inclusivity is not just ethical — it is a market opportunity and a sovereignty requirement.

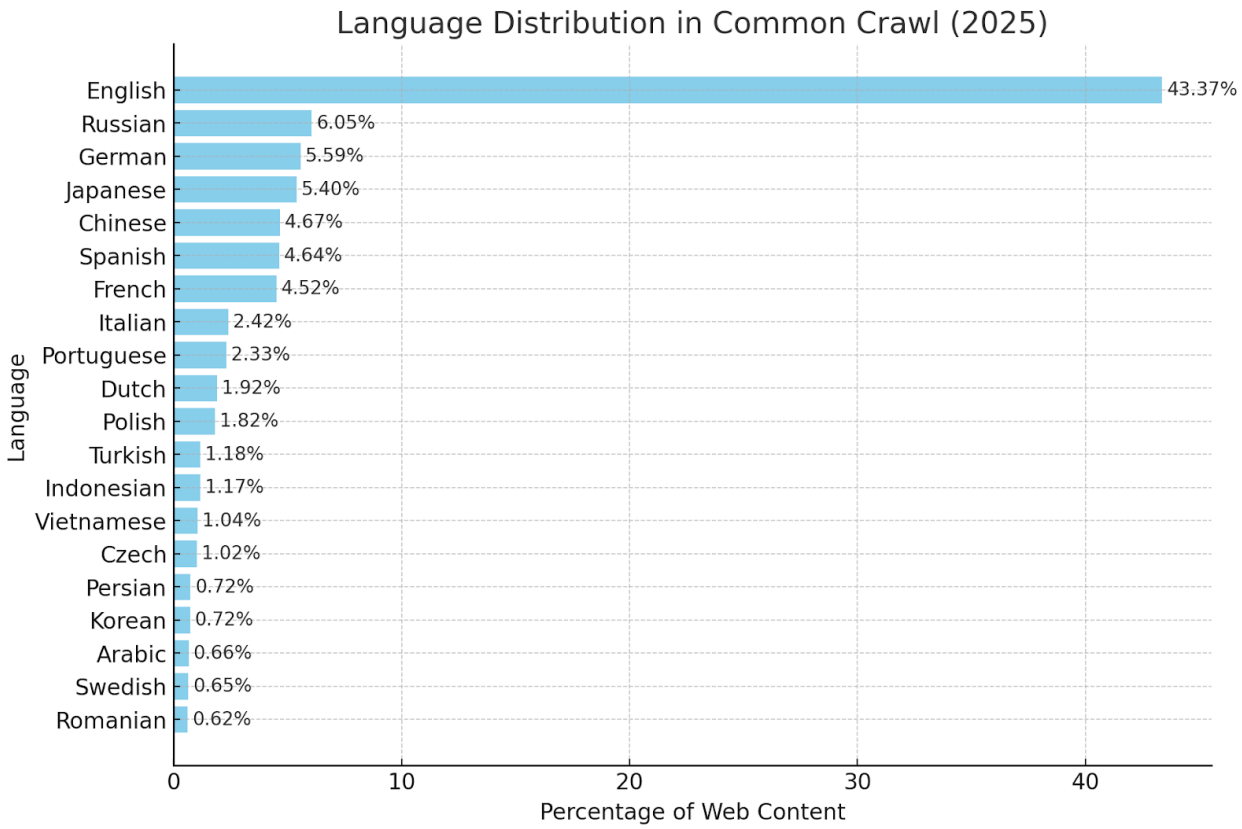
Omniscien™ TECHNOLOGIES

Global Population Distribution

There are more people inside the yellow circle than outside.

Omniscien™ TECHNOLOGIES

LLM Models are Biased



Language	Total Speakers	First Language Speakers	Second Language Speakers	Common Crawl Web Content %
English	1,456,448,320	375,000,000	1,081,448,320	43.37%
Mandarin Chinese	1,138,222,350	918,000,000	220,222,350	4.67%
Hindi	609,454,770	340,000,000	269,454,770	0.20%
Spanish	559,078,890	460,000,000	99,078,890	4.64%
Arabic	376,436,230	274,000,000	102,436,230	0.66%
French	309,804,220	77,200,000	232,604,220	4.52%
Bengali	272,828,760	230,000,000	42,828,760	0.11%
Portuguese	263,638,850	221,000,000	42,638,850	2.33%
Russian	254,997,130	154,000,000	100,997,130	6.05%
Urdu	231,717,940	70,600,000	161,117,940	0.00%
Indonesian	199,113,300	43,600,000	155,513,300	1.17%
German	133,245,880	76,100,000	57,145,880	5.59%
Japanese	123,445,570	122,000,000	1,445,570	5.40%
Korean	122,940,540	77,200,000	45,740,540	0.72%
Nigerian Pidgin	120,650,000	**	120,650,000	N/A
Cantonese	106,133,370	84,500,000	21,633,370	N/A
Marathi	100,016,870	83,200,000	16,816,870	0.03%
Telugu	95,981,790	83,000,000	12,981,790	0.02%
Vietnamese	93,077,700	85,000,000	8,077,700	1.04%
Wu Chinese	92,821,190	83,400,000	9,421,190	N/A
Turkish	90,028,000	75,700,000	14,328,000	1.18%
Tamil	86,640,030	75,000,000	11,640,030	0.04%
Tagalog	83,034,910	28,200,000	54,834,910	0.01%
Farsi	78,623,350	70,000,000	8,623,350	0.73%

Omniscien™ TECHNOLOGIES

A Clearer View of Language Representation

North America
385 million
(4.7%)

South America
663 million
(8.1%)

Europe
744 million
(9.0%)

Africa
1.5 billion
(18.50%)

Middle East
549 million
(6.7%)

Asia
4.8 billion
(58.7%)

Oceania
46 million
(0.6%)

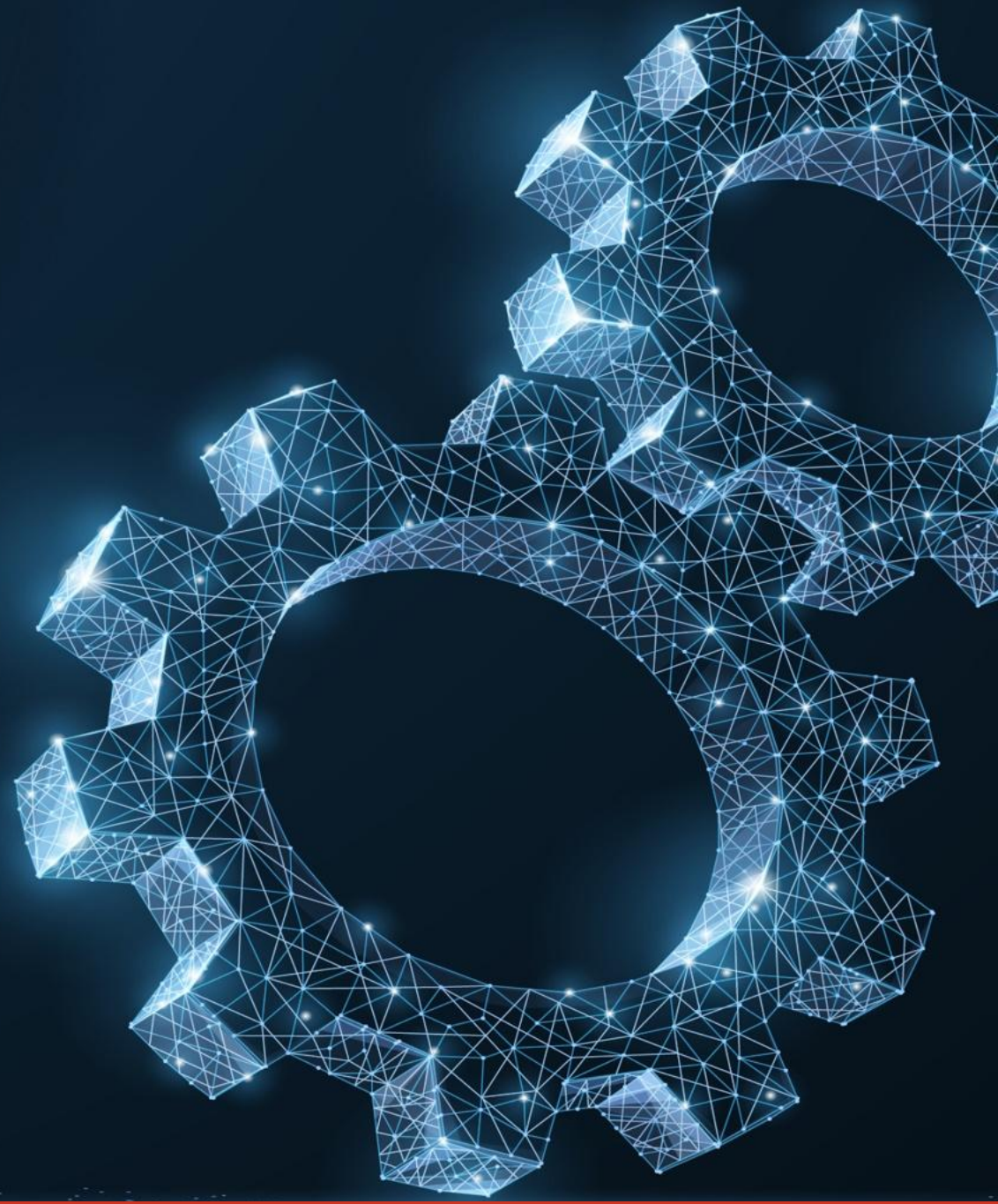
Omniscien™ TECHNOLOGIES

Omniscien™ TECHNOLOGIES

An Introduction to People-First AI

and

Why Every Organization Should Care About Digital Sovereignty



The Problem

Trust is Broken!

- The AI honeymoon is over
- AI hype has failed to deliver on its promises
- Confidence in AI—especially unethical Big AI vendors—has collapsed
- 99% of all U.S. companies have failed to scale generative AI successfully (McKinsey Global Survey)

Geopolitical + Compliance Pressure

- Geopolitical trust has been shattered with new unprecedented risks
- Regulatory and compliance pressure is rising
- Data privacy and sovereignty are non-negotiable

The AI hype phase is over. Now results matter.

- Perpetual POC purgatory—driven by 'me-too' AI boardroom mandates
- **Control** is now a primary buying driver

The cost? Millions wasted on failed AI pilots, compliance risks, and vendor lock-in—because without control, there is no scale.

Omniscien™ TECHNOLOGIES

We've moved beyond hype. Now it's about ***control, clarity, and trust****—delivered via* ***People-First and Sovereign AI***

SOLUTION

Language Studio Delivers a Two-Pronged Solution

People-First AI

Embeds intelligence into complex, repeatable human work, **automating what drains time, not talent.**

It doesn't chase enterprise-scale transformation. It focuses on **individual human processes, one at a time.**

Hundreds of small, **repeatable wins deliver more value, and rebuild more trust**—than any top-down AI rollout.

So that humans can move faster, think deeper, and lead with clarity, creativity, and control—**delivering the impact only people can**.

+

Digital Sovereignty

Ensures that nations, organizations, and individuals **retain full control over their data, infrastructure, and AI systems**—free from foreign interference, platform lock-in, or opaque dependencies.

It means the ability to deploy, govern, and scale AI on your own terms: **Where** you choose, **how** you choose, with **no external gatekeepers**.

Because **true trust and resilience** demand not just intelligent systems—but **systems you own, govern, and trust.**

People get clarity. Enterprises and nations get sovereignty. Everyone gets AI that works.

SOLUTION

People-First AI

Automate 70% of daily work

Removes the grind and mundane, redirects human time and expertise to work that delivers the most strategic value.

Initial Target Roles

- Sales and Marketing
- Procurement and Finance
- Legal and Compliance
- Research and Analysis
- Ops and Customer Support

Omniscien™ TECHNOLOGIES

Empowering People Through Automation. ***Driving Efficiency with AI-Powered Language Technology.***

80% of AI Projects Fail

The Failure Formula

- Buying AI ≠ transformation
- 81% of enterprise AI tools go unused after purchase *(IBM)*
- 85% of AI projects fail to deliver business value *(Gartner)*
- Most orgs invest 90% in tech, 10% in training

The result?

“Perpetual POC purgatory”

The problem isn’t AI—it’s how it’s bought, sold, and implemented.

They start with the tool, not the workflow.

They ignore the humans doing the work.

SOLUTION

Everyone’s buying AI. No one’s asking what it actually does.

The People-First AI Playbook

Don’t replace humans. Amplify them!!!

- Language Studio delivers value from day one—by designing around how people actually work and providing hundreds of pre-built templates
- Focus on **real workflows**, **immediate benefit**, not hype features
- Multiple early small-midsize wins build trust, momentum, and proof
- The 65/35 Rule:
  - 65% = process redesign, change management, workforce enablement
  - 35% = technology and infrastructure
- Even when the board insists to “just do AI,” we give the business a **real roadmap to success**

AI fails when it’s just a tool. ***It works when it’s part of the team.***

So how do we make AI actually work in the real world?

language STUDIO

People-First AI

SOLUTION

Reclaim Human Time

- **Built to Remove the Grind**

  Hundreds of prebuilt editable visual workflows eliminate high-effort, low-value admin tasks.

- **70% of Time, Reclaimed**

  Daily overhead gets automated—CRM updates, follow-ups, meeting notes, proposal drafts, scheduling, task routing, and more. Workflows flex to your business logic—**no rigid forms, no code.**

Scale What Humans Can't

, and more.

- **People-First AI Agents**

  Handle valuable-but-skipped tasks at scale—deeper insights, faster outcomes, no overhead.

- **Custom AI Templates**

  For board prep, legal review, doc analysis, interview logs, and more—trained on actual workflows, not toy demos.

Keep the Human in the Loop

- **Human-AI Collaboration**

  AI becomes a cognitive thinking partner—amplifying judgment, not replacing it.

- **Agentic AI**

  Plans and executes under full control, auditable, transparent—autonomy without risk.

Forget moonshots. Transformation starts with one real win — then another. Then another.

SOLUTION

Digital Sovereignty

Five Non-Negotiable Rights for Nations, Enterprises, and Critical Infrastructure

Model Autonomy
Build, deploy, and audit

Infrastructure Control
Where and how you compute

Data Control
Own, store, and govern your data and how it is used

Narrative Freedom
Control what's seen, said, and shared

Choice/Right to Exit
Freedom to leave, adapt, or self-host

5 Challenges Solved by One Platform

language STUDIO

Omniscien™ TECHNOLOGIES

AI Power Without Foreign Gatekeepers *Sovereign AI. No Permission Needed - Your Data. Your Models. Your Terms.*

The Five Pillars of Digital Sovereignty

SOLUTION

Infrastructure Control

Where and How You Compute

- Sovereignty is impossible if your compute lives in or is controlled by another jurisdiction.
- Control over infrastructure means the ability to deploy, scale, and operate independently of foreign policy, export controls, or revocable cloud APIs and “off-switches”.

Data Control

Who Holds It, Who Governs It, and Who Can Shut It Off

- If your data can be subpoenaed by a foreign court or throttled by a third-party provider, it’s not yours.
- Sovereignty demands lifecycle control—collection, storage, access, deletion

Model Autonomy

Who Builds, Trains, and Controls the AI You Depend On

- Most “AI platforms” rent you a black box. If you can’t retrain, self-host, or audit it, you’re not in control.

How Omniscien Addresses It:

- Fully deployable **on-premise or within sovereign cloud environments**.
- Supports **air-gapped, offline, and hybrid infrastructure** setups.
- Built with **open standards** and minimal external dependencies to prevent lock-in.
- **No hidden SaaS tethering**—you control the infrastructure, from chips to stack.

- **Full data residency control**—choose where data is stored, processed, and backed up.
- **Built-in data governance policies:** retention rules, audit trails, encryption, and role-based access.
- Data never leaves your jurisdiction unless you explicitly move it—**no silent syncs or external storage.**
- Fully compliant with **GDPR, ISO 27001, and regional data localization frameworks.**

- All models in **Language Studio** are **trainable, tunable, and auditable locally.**
- Provides **custom model training pipelines** for translation, transcription, classification, LLM workflows, and more—on your own data, in your own compute.
- **Supports open-weight models and custom inference stacks** for maximum independence.
- No forced API usage or “phone-home” behavior—**models are yours to inspect/modify.**

Omniscien™ TECHNOLOGIES

AI Power Without Foreign Gatekeepers

Sovereign AI. No Permission Needed - Your Data. Your Models. Your Terms.

The Five Pillars of Digital Sovereignty

SOLUTION

Narrative Freedom

Who Defines What Can Be Said, Seen, or Even Asked

- Narrative sovereignty is erased when platform bias, moderation filters, or training data exclusions distort how your culture, language, or discourse is represented.

Choice/Right to Exit

Exit, Adapt, or Build on Your Own Terms

- If you can’t leave, fork, or migrate, then you're not sovereign—you’re trapped.

How Omniscien Addresses It:

- **Supports low-resource, minority, and sovereign languages** with dedicated pipelines.
- All NLP and translation tools are **language-customizable**, including dialect and terminology management.
- **No externally dictated prompt filtering or censorship logic**—you define moderation boundaries.
- Enables you to **train models on local corpus, cultural content, and policy-aligned guidelines.**
- Entire platform is **modular, portable, and self-hostable**—no dependency on Omniscien-operated services.
- **Contracts support subscription licenses**, local installations, and full data/model export.
- Built with interoperable open standards—easy to integrate, fork, or migrate without losing capability.
- Glue for open and closed, proprietary and integration blah blah
- **You own the system logic**, not just the interface —your stack, your control.

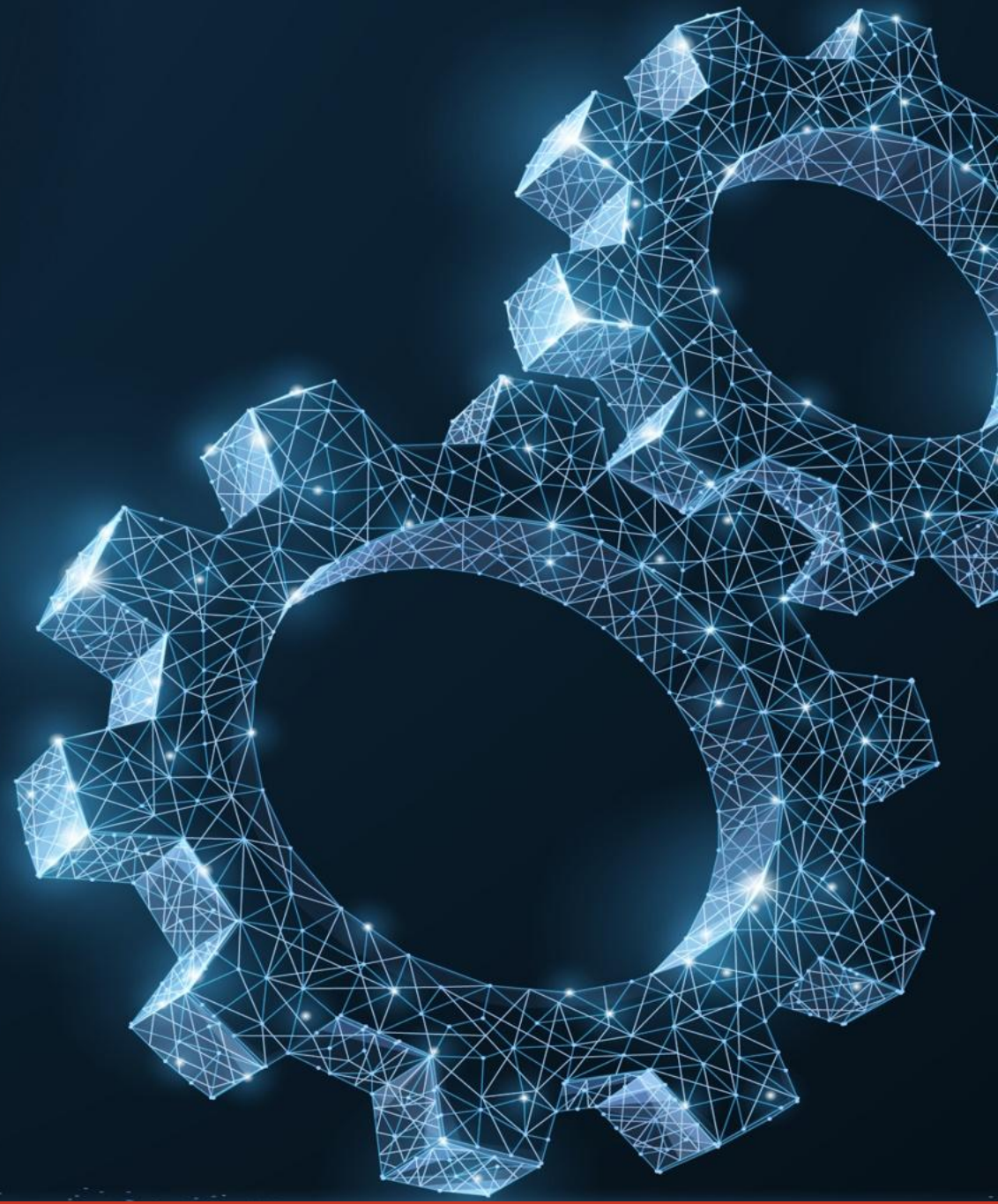
Omniscien™ TECHNOLOGIES

AI Power Without Foreign Gatekeepers

Sovereign AI. No Permission Needed - Your Data. Your Models. Your Terms.

Omniscien™ TECHNOLOGIES

Official Launch of Language Studio 7



Beyond the Hype

Omniscien™ TECHNOLOGIES

language STUDIO

We built the sovereign platform that the post-hype AI world demands — and we're already deploying it.

Control, Compliance, Sovereignty – and Rapid Success.

AI's next era won't be hype-driven. It will be sovereignty and outcome-driven. That's us.

language STUDIO

SOLUTION

Translate

650+

Machine Translation (MT) Language Pairs

Transcribe

55

Automatic Speech Recognition (ASR) Languages

Recognize

210+

Optical Character Recognition (OCR) Languages

Convert

200+

File and Document Conversion Formats

Analyze

200+

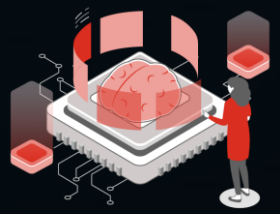
Document and Natural Language Processing (NLP) Tools

Conversational AI

55

Conversational Languages to Engage in Meaningful Dialog

Business AI



100+

Powerful AI Tools for Everyday Business Use Cases

Linguists

70+

Linguistic Analysis and Quality Control Tools

Workflows

50+

Pre-Build Workflow Templates for Business

Broadcast

60+

Advanced Media Processing (AMP) and Broadcast Specific Tools

Generative AI

800K+

AI Models Supported for Customization and Fine-Tuning

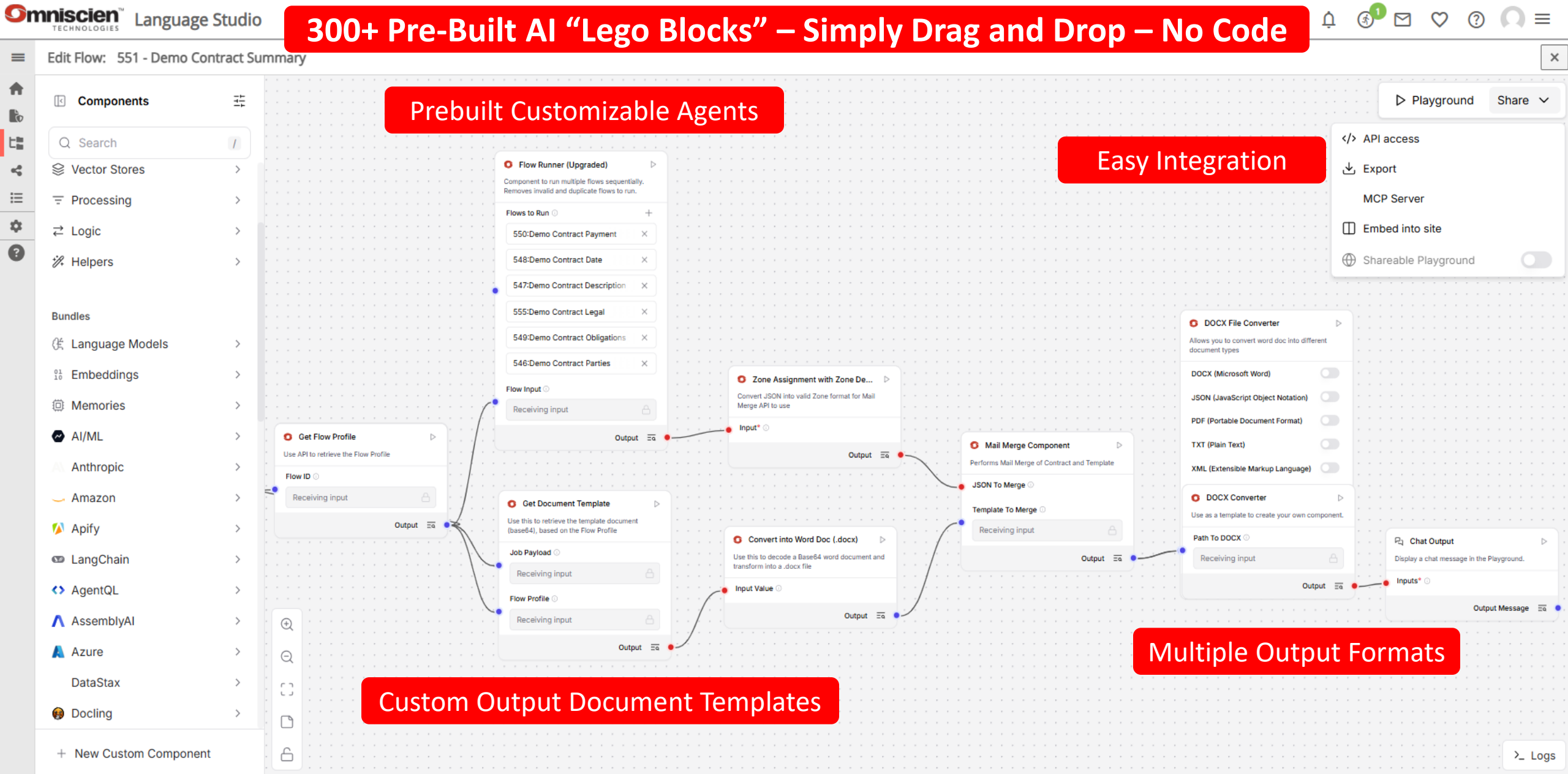
Text to Speech

140

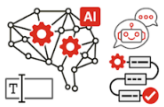
Languages Spoken with hundreds of voices and voice cloning.

Modular by Design. Built for Complex, Language-Centric Workflows. Unified by Purpose for Enterprise AI at Scale.

300+ Pre-Built AI "Lego Blocks" – Simply Drag and Drop – No Code



- Home
- Secure Tools
  - Business AI
  - Convert
  - Data Privacy
  - Linguist
  - Media
  - Office
  - Recognize
  - Transcribe
  - Translate
  - Manage Profiles
- Project Tools
  - Organizations
  - Projects
  - Agents
  - Classifiers
  - Generative AI
  - Models
  - Prompts
  - Workflows
- File Share
- Jobs
- Settings
  - Favorites
  - My Profile
- Help
  - API Information
  - Documentation
  - Technical Support



Business AI

Streamline your data analysis, summarization, and research with cutting-edge Generative AI tools.

Harness powerful AI to automate tasks and enhance efficiency across documents, meetings, forms, and custom chatbots, all with seamless integration and intelligent prompts that redefine productivity.

Audio, Video, Image, and Document Analysis

Analyze a variety of media formats, including audio, video, images, and documents, to extract essential information, generate summaries, and standardize outputs. Whether working with meeting recordings, video content, or text data, these tool streamlines analysis, saving time and ensuring accuracy across all formats.

Analyze My Documents
Generate concise summaries and formatted outputs from your documents with ease. Using predefined templates, this tool analyzes reports, contracts, and other document types, extracting key information to save time and ensure consistency across all content.

Analyze My Media
Utilize advanced AI to analyze video content, detect objects, and perform sound analysis. This tool identifies key elements such as objects, scenes, and audio patterns, offering detailed insights and summaries that streamline media processing and enhance content understanding.

Analyze My Meeting
Analyze meetings from video, audio, and calendar data to produce comprehensive summaries, minutes, action items, discussion points, and speaker analyses using predefined templates.

Normalize My Documents
Normalize various documents such as resumes, contracts, meeting minutes, and invoices into standardized formats using predefined templates.

AI-Driven Communication and Interaction

Leverage secure AI-driven tools for seamless communication and interaction, from building chatbots to querying documents, empowering smarter and faster problem-solving.

Ask My Documents
Upload a collection of documents (PDF, Word, PowerPoint, EPub, or Text formats) and utilize Generative AI to ask questions about specific documents or across the entire document set.

Build My Chatbot
Create a custom chatbot tailored to your needs: integrate it into your website for customer support, lead generation, and user engagement. Easily upload data or crawl web content, then embed it with a simple HTML code or API call.

Generative AI Chat
Chat directly with Large Language Models (LLMs) or use predefined workflows and prompts. It feels like ChatGPT but with one major difference: it's completely secure and private. Ask questions, get insights, or perform tasks efficiently with AI models tailored for safe and versatile communication.

Navigate My Data
Interact with your data using AI-powered navigation tools. Conduct detailed searches, link various information sources, and efficiently locate insights across different datasets. Navigate through complex data while streamlining the research process to identify key patterns, connections, and trends that support data-driven decisions.

Content Authoring and Refinement

Enhance and refine your written content with AI-powered tools that ensure consistency in style, vocabulary, and terminology, while generating new content and auto-populating forms.

Analyze My Content
Ensure consistency and precision in your written content. This tool analyzes text for style, terminology, vocabulary, and locale alignment, providing detailed reports, edited versions, and actionable suggestions for improvement. Perfect for maintaining uniformity in documents, blogs, emails, and more.

Answer My Form
Analyze forms in Adobe PDF, Microsoft Word, or Excel formats and auto-populate them with information extracted from provided reference documents.

Generate My Image
Create custom images with AI tailored to your needs. Simply provide a description or reference materials, and this tool generates high-quality visuals for presentations, reports, or marketing materials. Customize the style, format, and details to ensure the image aligns with your project's requirements.

Write My Content
Craft blog posts, emails, articles, press releases, customer responses, and more by referencing uploaded content for accurate context and detailed information.

Write My Presentation
Generate well-structured presentations with the help of AI. Provide your key points or upload reference materials, and this tool will craft slides that clearly convey your message. Automatically format content, suggest visuals, and ensure a logical flow, saving time and ensuring consistency in your presentations.

AI-Powered Automation

Automate repetitive tasks and workflows with advanced AI-driven solutions, improving efficiency and accuracy while allowing you to focus on high-priority tasks.

Automate My Desktop
Simplify and automate desktop tasks effortlessly. Record mouse clicks and keystrokes to create custom workflows that perform repetitive actions across your desktop applications. Boost productivity and minimize manual effort with powerful automation, from data entry to file management.

Automate My Processes
Easily automate routine workflows, from document analysis to form population and data standardization. Leverage advanced AI to streamline tasks, save time, and ensure accuracy with minimal manual effort. Let AI handle repetitive tasks, allowing you to focus on what matters most.

Batch Prompt Runner
Sequentially execute multiple predefined prompts from a user-created batch file.

Data Visualization

Advanced Media Processing (AMP)

Essential tools for organizations that require professional real-time and batch mode audio, video, subtitle, and caption processing.

Features include high-quality transcription, translation, subtitle adjustments and refinement, dialog extraction and alignment to timecodes, terminology and entity extraction, directors script extraction and creation, water marks, auto-conforming and re-conforming subtitles, automated language ID, metric tools, and more.

Transcribe

A comprehensive set of tools for converting audio and video into accurate, readable text. Enhance transcription quality, transcribe and translate audio or video files, and generate various output formats including subtitles, captions, draft dialogs, and formatted documents. Support live transcription of audio/video streams and interviews with real-time translation options, ensuring precise and accessible transcriptions for various applications and record keeping.

Analyze Transcriptions
Analyzes characters, chapters, scenes, situation, tone, key elements and notable vocabulary to assist human linguists and other automated processes.

Automated Transcription Refinement
Enhance transcription quality by improving readability, fluency, naturalness, and grammar. Adjust sentence boundaries, tailor for locale, remove utterances, and autocorrect misheard words.

Human Transcription Services
Add humans in the loop for post-editing and quality assurance of transcriptions and machine-translated content, or use humans for all transcription, translation, editing, and quality assurance processes.

Transcribe Audio/Video File
Transcribes and translates audio or video files to Voice JSON and a range of output formats including subtitles, captions, draft dialogs, data files and formatted Microsoft Word documents.

Transcribe Draft Dialog
The draft dialog can be easily reviewed, with any mishears corrected and adjusted. The reviewed dialog can then be processed through the Time/Dialog Alignment tool.

Transcribe Live Audio/Video Stream
Transcribes and translates live audio or video streams into formatted subtitles/captions, customized Microsoft Word transcripts and scripts, draft dialogs for human review, data files for use in other applications, and text transcriptions.

Transcribe Live Interview
Transcribes a live interview between 2 people using 2 microphones with optional real-time translation. Generate custom formatted Microsoft Word transcripts, subtitles, and data files for use in other applications and record keeping.

Translate

A robust suite of tools designed to translate various forms of media content accurately and efficiently. Enhance subtitle translations with AI-powered post-editing, translate subtitles and captions between languages while maintaining formatting, and automatically translate video synopses. Additionally, translate user reviews, capturing the unique style and colloquialisms to produce natural-sounding reviews in different languages.

Automated Post Editing
Enhance subtitle translation quality with AI-powered post-editing, which automatically refines both human and machine translations based on preset preferences, enforcing glossaries, writing styles, character adaptation, and more.

Human Translation Services
Add humans in the loop for post-editing and quality assurance of machine-translated content, or use humans for all translation, editing, and quality assurance processes.

Stylize To Speaker Character Profile
Adapt transcriptions and translations to match the unique speaking style of each speaker, ensuring the text reflects their true character and manner of speaking.

Translate Subtitles/Captions
Translate subtitles/captions from one language to another while adhering to a preconfigured formatting profile to produce high-quality structured subtitles/captions in the target language.

Translate Synopsis
Automatically translate video synopsis with machine translation engines specifically optimized for purpose.

Translate User Reviews
User reviews have their own unique style that includes slang and colloquialisms. Translate within the context of movies producing natural-sounding reviews across languages.

Conforming and Alignment Tools

A specialized set of tools for ensuring precise synchronization and formatting of subtitles and captions. Automatically conform subtitles to match audio modifications, re-conform subtitle files to adhere to preconfigured formats and structures, and align dialog files with time cues from audio or video transcriptions. Generate various output formats, including subtitles, captions, draft dialogs, data files, and formatted documents, ensuring accuracy and consistency in your media content.

Auto-Conform Subtitles/Captions
Automatic analysis of audio to identify modifications such as deletions or additions (e.g., censored, trimmed or substituted content) and adjustments to the provided time syncronized subtitle/caption file.

Re-Conform Subtitles/Captions
Automatically re-conform a subtitle/caption file to match preconfigured formats, structure, and features. Adjust various elements such as reading speed, frame rates, and characters per line; rectify improper formats; and convert between subtitle and caption formats.

Time/Dialog Alignment
Automatically synchronize and align a dialog file to time cues generated from an audio or video transcription. Generate a range of output formats including subtitles, captions, draft dialogs, data files and formatted Microsoft Word documents and as-spoken directors scripts.

Video and Audio Tools

A versatile collection of tools designed to enhance and manipulate video and audio content. Add text or image watermarks to videos, capture video images at specific timestamps, and compress videos for optimal file size. Separate audio stems to extract individual components like vocals and instruments, and generate detailed descriptions for both video and image content through advanced analysis.

Add Watermark to Video

Analyze Voice Biometrics

Translate

Translate documents, data files, and text between languages using the latest state-of-the-art neural machine translation technology. Text and HTML are translated instantly. Files and documents are batch processed as jobs which can be downloaded on completion.

Common Translation Tools

A suite of essential translation tools, offering quick access to the most commonly used features for efficient and accurate translations.

Quick Translate (Text and HTML)
Translate instantly by typing or copy and pasting text or HTML into a form.

Translate Files
Translate documents and files from a variety of formats. Simply drag and drop your files to translate.

Translate Adobe PDF Documents
Use OCR to convert Adobe PDF documents and any images contained within to Microsoft Office formats, then translate the file to another language.

Translate Images
Use OCR to convert images to Microsoft Office formats, then translate the file to another language.

Show my Translation Jobs
Show the history of files that you submitted to translate and download recently processed file jobs.

Basic File Translation Tools

Translation Tools offer a range of features including the ability to translate various file formats such as PDFs, DITA, HTML, images, JSON, Microsoft Office documents, presentations, spreadsheets, text files, voice JSON, and XML, as well as view and download the history of translation jobs.

Translate DITA Files
Translate DITA XML files from one language to another.

Translate HTML Files
Translate HTML files from one language to another, retaining formatting.

Translate JSON Files
Translate JSON files from one language to another, with optional processing rules.

Translate Microsoft Word
Translate Microsoft Word (DOC/DOCX) documents from one language to another keeping the original formatting.

Translate Presentations
Translate Microsoft PowerPoint (PPT/PPTX) presentations from one language to another keeping the original formatting.

Translate Spreadsheets
Translate Microsoft Excel (XLS/XLSX) spreadsheets from one language to another. See Special Translation Tools for more complex features

Translate Text Files
Translate text files from one language to another.

Translate Voice JSON Files
Translate Omniscien Voice JSON data files from one language to another.

Translate XML Files
Translate XML files from one language to another, with optional processing rules.

Advanced Translation Tools

Advanced Translation Tools provide targeted solutions for complex translation tasks, including managing multilingual profiles to guide translation settings, translating diverse file formats to English, translating documents that include more than one language, locating and translating specific parts of Excel sheets, translating subtitles with automatic format adjustments, translating emails with attachments, and handling industry-standard TMX and XLIFF files.

Translate Any Document to English
Detect the file format and the source language and translate the content to English.

Translate Parts of an Excel Spreadsheet
Locate sheets and columns using Regular Expressions, automatically detect the language of each cell in a column, and translate to another language.

Translate Subtitles
Translate subtitles (SRT/TTML/DFXP/PAC) from one language to another. Automatically adjust ouput format and structure based on rules.

Translate Emails
Translate emails from multiple languages into a single language, including their attachments.

Translate Multilingual Documents
Automatically detect multiple languages within document and converts them all into a single target language, ensuring a seamless and unified translation

Translate Translation Memories
Translate industry standard TMX and XLIFF files from one language to another.

Translate and Analyze Documents
Translate then use Business AI tools to analyze document nt contents.

Translate and Normalize Documents
Translate then use Business AI tools to normalize ze document formats.

Home
Secure Tools
Business AI
Convert
Data Privacy
Linguist
Media
Office
Recognize
Transcribe
Translate
Manage Profiles
Project Tools
Organizations
Projects
Agents
Classifiers
Generative AI
Models
Prompts
Workflows
File Share
Jobs
Settings
Favorites
My Profile
Help
API Information
Documentation
Technical Support

Analyze My Meeting

Analyze and report on various meetings, including All-Hands, Board, Budget Planning, Client, Sales Calls, and more.
Gain insights and actionable recommendations across multiple categories such as meeting analysis, strategic planning, compliance, and crisis management with detailed analyses tailored to each meeting type, providing comprehensive summaries, key highlights, action items, and sentiment assessments for informed decision-making.

Meeting | Document Groups | Jobs (0, 0)

Search | Show filters

Board Meeting

Summarize key strategic decisions and action plans. Highlight and summarize financial discussions and projections. Identify discussions around compliance and governance issues, and gauge the sentiment of board members on critical issues to understand consensus and concerns.

Meeting Analysis 246

Brainstorming Session

Group similar ideas to identify key themes. Measure the diversity and novelty of ideas to foster creativity. Highlight actionable ideas and next steps to move from ideation to implementation, and track participation levels to identify active contributor

Meeting Analysis 247

Budget Planning Meeting

Summarize discussions on budget planning and allocation. Highlight key financial priorities and constraints. Identify potential risks and opportunities, and track action items and responsibilities for budget preparation.

Meeting Analysis 274

Budget Review Meeting

Analyze financial performance against the budget. Summarize discussions on budget adjustments and allocations. Highlight key financial concerns and recommendations, and track decisions on resource allocation.

Meeting Analysis 253

Client Meeting

Summarize feedback and suggestions from the client. Identify and prioritize next steps agreed upon. Assess the sentiment and satisfaction level of the client, and analyze the overall health of the client relationship to ensure long-term collaboration.

Meeting Analysis 249

Community Outreach Meeting

Summarize discussions on community engagement and initiatives. Identify and categorize community concerns and suggestions. Track progress on community projects and events, and measure overall sentiment and engagement.

Meeting Analysis 275

Compliance Meeting

Summarize discussions on compliance requirements and issues. Identify potential compliance risks and mitigation strategies. Track progress on compliance initiatives, and measure overall adherence to regulatory standards.

Meeting Analysis 270

Crisis Management Meeting

Identify and summarize key issues and challenges discussed. Track decisions and action plans for crisis resolution. Analyze potential impacts and mitigation strategies, and measure overall sentiment and readiness to handle the crisis.

Meeting Analysis 255

Cross-Functional Meeting

Summarize discussions involving multiple departments or teams. Identify and categorize collaboration opportunities and challenges. Track action items and responsibilities across teams, and measure overall alignment and cooperation.

Meeting Analysis 269

Design Review Meeting

Summarize feedback on design proposals. Highlight key design decisions and rationale. Identify areas for improvement and potential design risks, and track action items for further refinement.

Meeting Analysis 261

Employee Onboarding Meeting

Summarize key points and objectives of the onboarding process. Identify areas where new hires may need additional support or information. Track progress on onboarding tasks and training, and measure overall satisfaction and readiness of new employees.

Meeting Analysis 268

General Online Meeting

Automatically analyze and summarize online meetings, capturing key points, action items, and participant insights. Compatible with platforms like Zoom, Teams, and Google Meet, it ensures no important detail is missed, streamlining follow-up with tailored summaries and action plans.

Meeting Analysis 517

Incident Review Meeting

Summarize the incident details and the impact. Identify root causes and contributing factors. Discuss and track action items for preventing future incidents, and measure overall readiness and response effectiveness.

Meeting Analysis 262

Innovation Meeting

Summarize discussions on new ideas and innovative solutions. Identify and categorize proposed innovations and their potential impact. Track action items for further exploration and development, and measure overall creativity and engagement.

Meeting Analysis 273

Leadership Meeting

Summarize key points and decisions from discussions among leadership. Highlight strategic goals and priorities. Identify potential challenges and opportunities, and track action items and responsibilities among leadership.

Meeting Analysis 271

Operational Meeting

Summarize discussions on operational efficiency and improvements. Identify key performance metrics and areas for improvement. Track action items for operational changes, and measure overall operational effectiveness.

Meeting Analysis 266

Partnership Meeting

Summarize key points from discussions with potential or current partners. Identify and categorize partnership opportunities and challenges. Track progress on partnership agreements and initiatives, and measure overall sentiment and satisfaction.

Meeting Analysis 265

Performance Review Meeting

Summarize feedback on individual or team performance. Track progress towards performance goals and objectives. Identify strengths and areas for improvement, and outline next steps and development plans.

Meeting Analysis 254



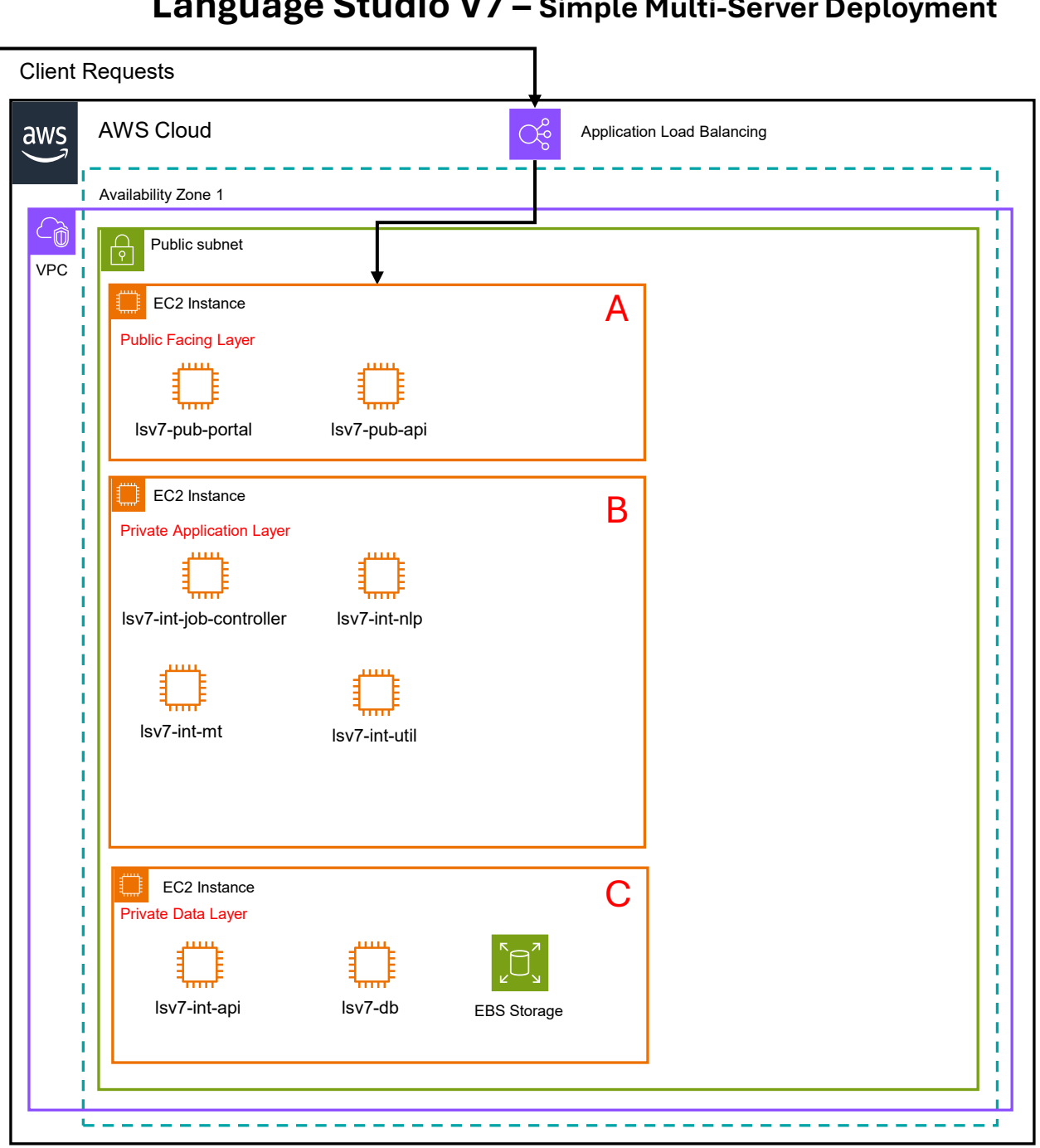


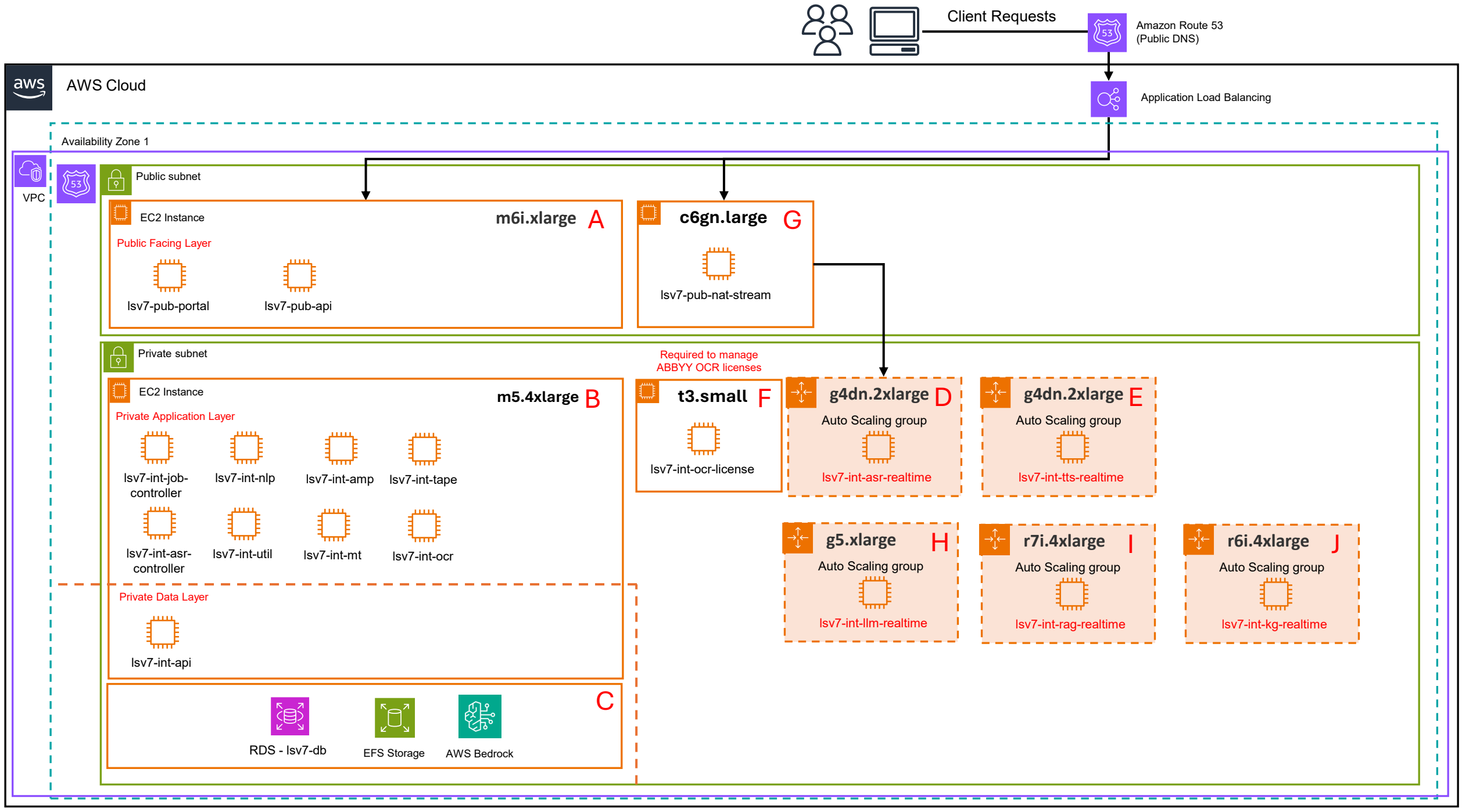




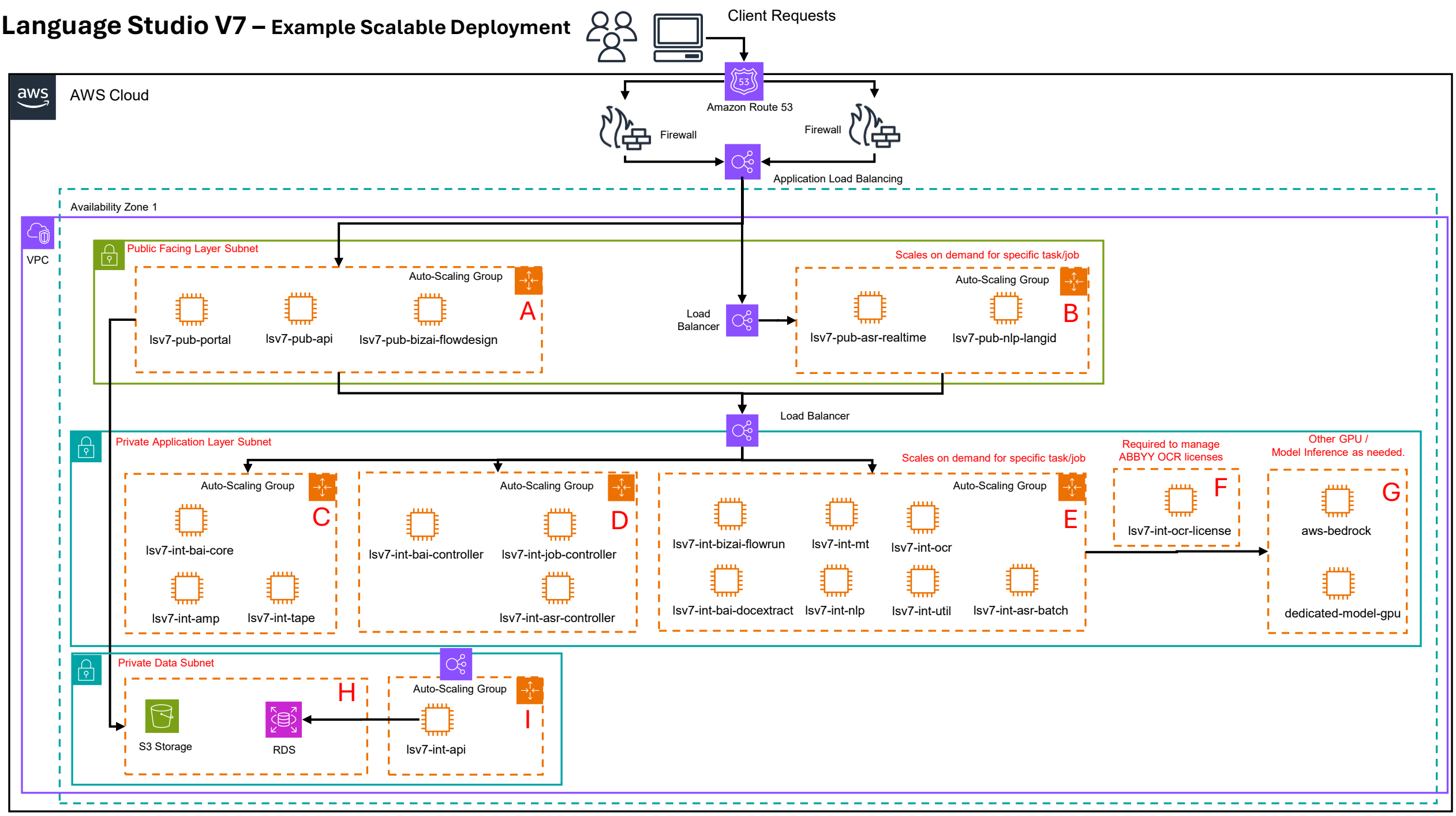
Logical/Physical Architecture

- **Single-server deployment with layered architecture:**
  - All components run on a single EC2 instance within a public subnet in AWS.
  - Exposed to clients via an Application Load Balancer.
- Three logical layers:
  - Public Application Layer:
    - Handles external web and API requests (lsv7-pub-portal, lsv7-pub-api).
  - Private Application Layer:
    - Manages core processing services such as NLP, MT, job control, and utilities (lsv7-int-job-controller, lsv7-int-nlp, lsv7-int-mt, lsv7-int-util).
  - Private Data Layer:
    - Contains internal APIs and database (lsv7-int-api, lsv7-db) with EBS-backed persistent storage.
- Deployed in a single Availability Zone and single VPC for simplicity and ease of maintenance.
- Suitable for low to moderate workloads or development/testing environments.
- This is a basic system only for simple NLP tasks and translation.
- Database is contained on the same server rather than using AWS RDS features. RDS could be easily substituted.




Client Requests
Amazon Route 53
(Public DNS)
AWS Cloud
Application Load Balancing
Availability Zone 1
VPC
Public subnet
EC2 Instance
m6i.xlarge A
Public Facing Layer
lsv7-pub-portal
lsv7-pub-api
c6gn.large G
lsv7-pub-nat-stream
Private subnet
Required to manage
ABBYY OCR licenses
EC2 Instance
m5.4xlarge B
Private Application Layer
lsv7-int-job-controller
lsv7-int-nlp
lsv7-int-amp
lsv7-int-tape
lsv7-int-asr-controller
lsv7-int-util
lsv7-int-mt
lsv7-int-ocr
Private Data Layer
lsv7-int-api
C
RDS - lsv7-db
EFS Storage
AWS Bedrock
t3.small F
lsv7-int-ocr-license
g4dn.2xlarge D
Auto Scaling group
lsv7-int-asr-realtime
g4dn.2xlarge E
Auto Scaling group
lsv7-int-tts-realtime
g5.xlarge H
Auto Scaling group
lsv7-int-llm-realtime
r7i.4xlarge I
Auto Scaling group
lsv7-int-rag-realtime
r6i.4xlarge J
Auto Scaling group
lsv7-int-kg-realtime

Language Studio V7 – Example Scalable Deployment



Our Moat: Full Stack Control, Compliance, and Sovereign-Ready People-First AI Infrastructure

Omniscien™ TECHNOLOGIES

Deep Structural Differentiation

- Fully self-hostable and modular platform
- Supports for air-gapped, offline, and hybrid deployments
- No dependency on SaaS APIs or vendor cloud platforms
- Built with open standards not locked behind proprietary walls

Compliance-First Architecture

- Native support for GDPR, ISO 27001, EU data residency laws
- Meets emerging AI Act localization and control mandates
- Enables AI deployments in regulated industries (defense, government, healthcare)
- Eliminates off-switch risk from foreign jurisdictions

Control-Centric Integration

- Deep integration with enterprise workflows (translation, transcription, LLMs)
- Custom model training pipelines for local data and local languages
- Customers own their models, data, and infrastructure logic
- Exit-ready. Can migrate, fork or expand without us

Pre-Built People-First AI

- Hundreds of pre-built workflows for horizontal specialties
- Horizontal workflows that eliminate high-effort, low-value work = **quick wins**
- Designed for strategic roles: legal, compliance, research, ops
- Augments judgment, decision-making, and oversight

We're not just a product. We're the infrastructure for every pillar of Sovereign AI

Omniscien™ TECHNOLOGIES

language STUDIO

Dion Wiggins
Founder & Chief Technology Officer,
Omniscien Technologies

Professor Philipp Koehn
Chief Scientist
Omniscien Technologies

Professor of Computer Science,
Johns Hopkins University

Reality Check: AI and Language Processing in 2025

What Works, What Fails, and What Matters Next