



Frontier AI. In your hands.

Configurable AI for Countries

Agenda

1. Mistral AI For Countries
2. Mistral Products & Solutions
3. Mistral Success Stories

We are Mistral AI.

Founded on deep scientific expertise.

 Arthur Mensch
Co-author of the original Gemini paper

Google

 Guillaume Lample
Co-author of the original Llama paper

Meta

 Timothee Lacroix
Co-author of the original Llama paper

Meta

+ 400-strong team

With a clear mission.

Empower customers to maximize the power of frontier AI and create meaningful value



Prioritizing real customer outcomes.

 Time-to-sustained value over superficial prototypes

 ROI over benchmarks

 Customer control, security, and customizability over race to AGI

Mistral AI is the global leader in helping public agencies transform with AI

The world's leading independent AI lab.

- Only non-aligned frontier AI lab
- Proudly independent, supported globally
 - Deep focus on cultural alignment, regional values and priorities

Vertically integrated.

- Local compute and datacenters
 - Wide choice of customizable frontier models
 - Full AI lifecycle management
- Custom, large-scale AI products and solutions

Open and local.

- World leader in open models
 - Data, weights, guardrails that remain in control of the customer
- Flexibility of local deployment, edge, or any cloud
 - Deeply customizable across the stack

Turnkey expertise at your service.

- Full service across use case development, applied AI, and large-scale deployments
- Complete transformation of public services from AI-first principles
- Proliferation of AI across public services
 - Deeply partnered R&D
- Continuous adaptation to evolving data, policies, and markets



A transformative AI solution should be centered around **your competencies**.

Your:

Transform into...

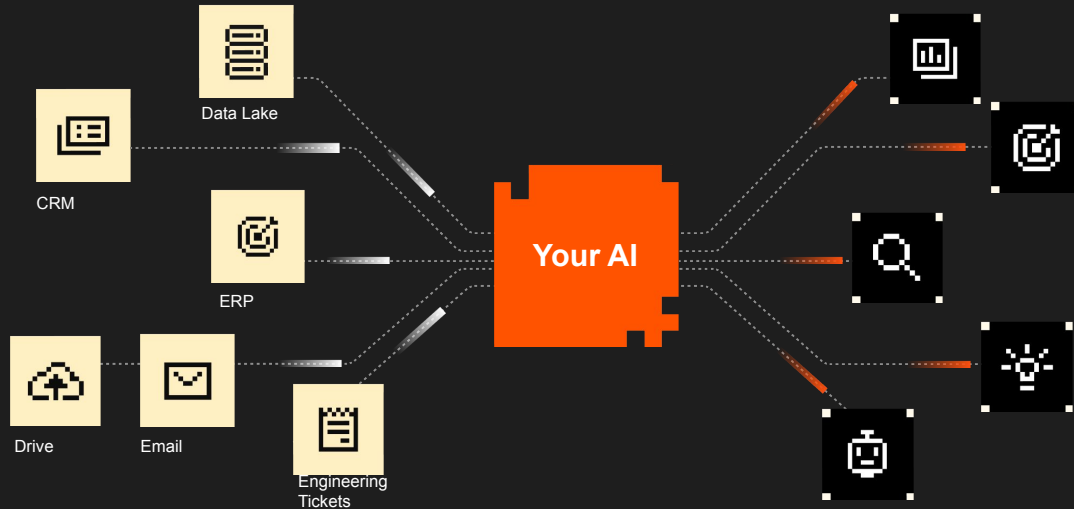
Your long-term competitive edge

Expertise

Processes

Data

Insights



A transformative AI solution turns your competencies into your **AI flywheel**.

When your:

Transform into...

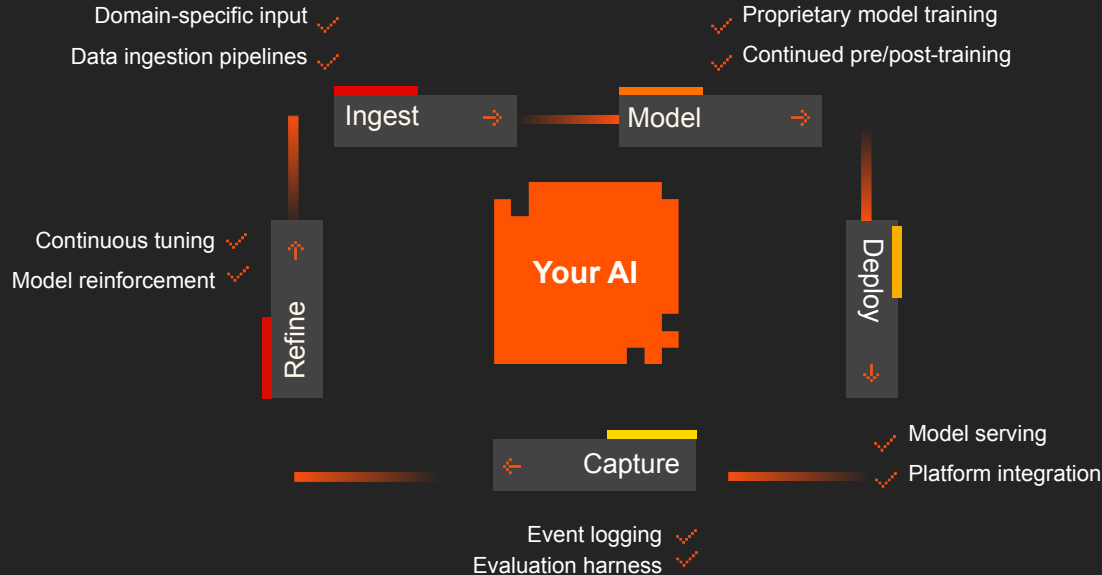
Your long-term competitive edge

Expertise

Processes

Data

Insights



Smarter decisions ✓

Autonomous operations ✓

Deeper insights ✓

Continuous improvement ✓

Driving AI Transformation for our customers

Internal enablement

Internal agents



Advisors support chatbot



Customer facing chatbot



Core Process Automation

Strategic Partnership



Legal Assistant



LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG

Smart Cargo Release



Customer & market innovation

In-car voice assistant



Co-trained SOTA LLM for SEA Languages



Co-training custom VLAs



And many other ...



BNP PARIBAS



MINDEF
SINGAPORE



Products

The Mistral AI **model landscape** - all available on-prem

Premier models



**Mistral
Medium 3**

Our **flagship** model, for your **most sophisticated** needs



**Magistral
Medium**

Our **frontier-class reasoning** model



Ministral 3B/8B

Our most **efficient / powerful edge** models



**Mistral
Moderation**

Our **moderation service** to detect **harmful** text content



Codestral

Our **low-latency code model** for code generation (FIM)



**Devstral
Medium**

Our **agentic LLM** for high-performance software engineering tasks



Mistral Embed

Our **embedding model** to make your data accessible



Mistral OCR

The world's **best OCR model** to-date

Open-source models *Apache 2.0*



**Voxstral
Small/Mini**

Our **frontier speech understanding** models



**Mistral
Small 3.2**

Our **enterprise-grade small model**, cost efficient and fast



**Magistral
Small**

Our **small reasoning** model

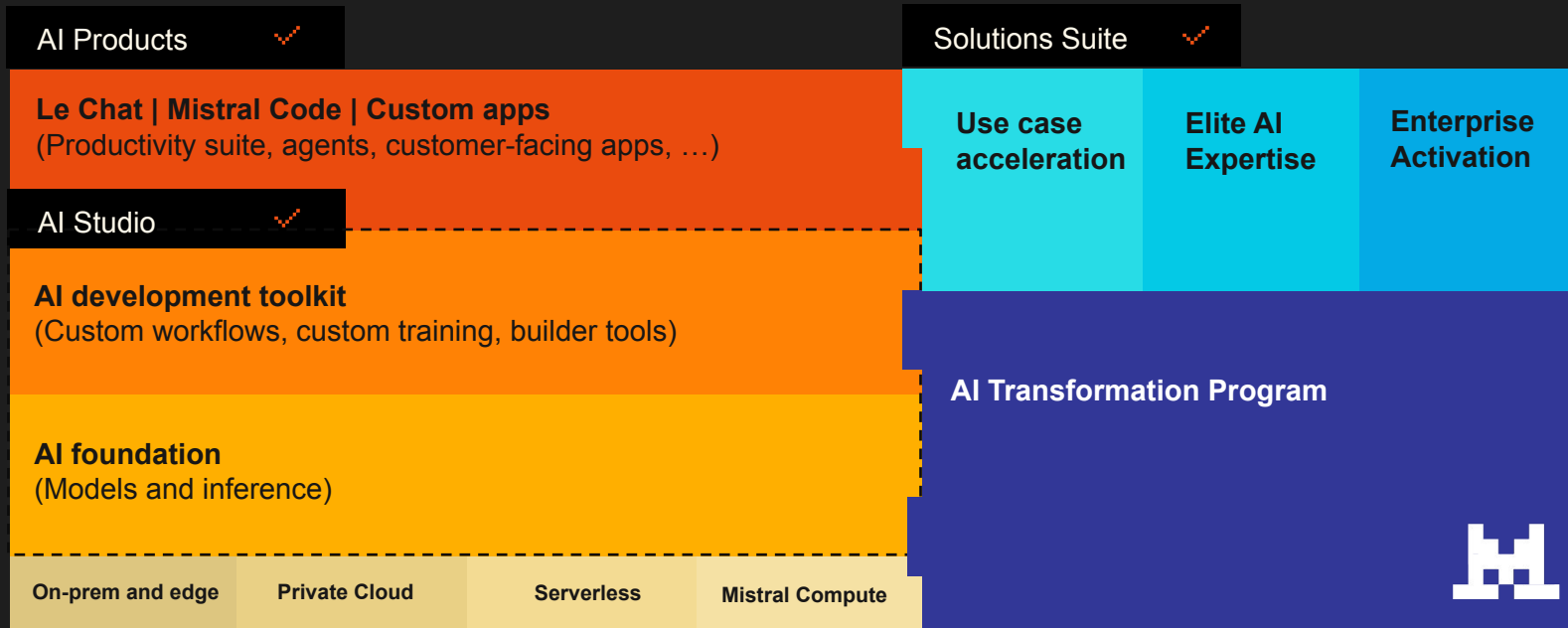


Devstral Small

Our **small agentic LLM** for software engineering tasks



The Mistral AI enterprise suite



Le Chat enterprise for individual productivity

Science at Work

Cutting-edge models, tailored to your needs

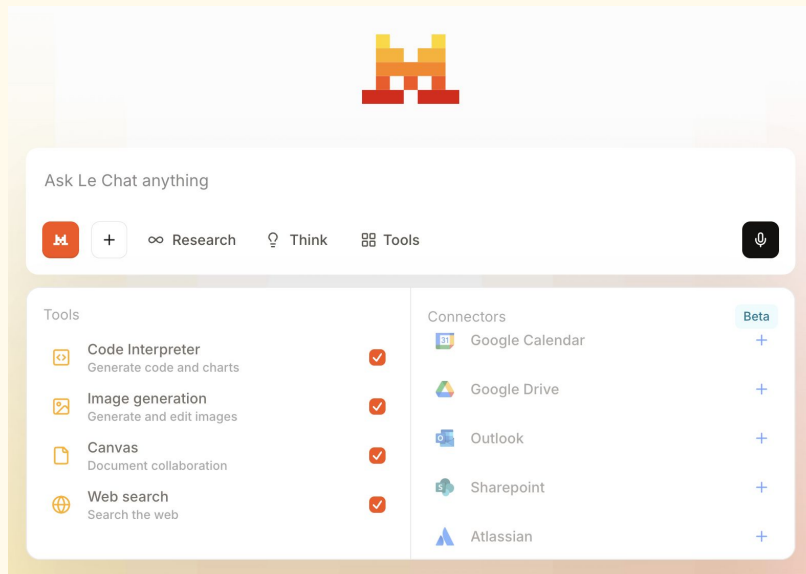
Access to the latest generation of models — from general-purpose to specialized reasoning and deep research — with the ability to create custom agents



Natively multi-functional

All the tools you need, built in

Code execution, web search, image generation, and Canvas collaboration — expanding the assistant's scope without leaving the chat



Fully integrated

Seamlessly connected to your ecosystem

Native connectors for Google (Gmail, Calendar), Microsoft (Outlook, SharePoint), Atlassian, and beyond — plus full MCP compatibility for any system



Personalized by design

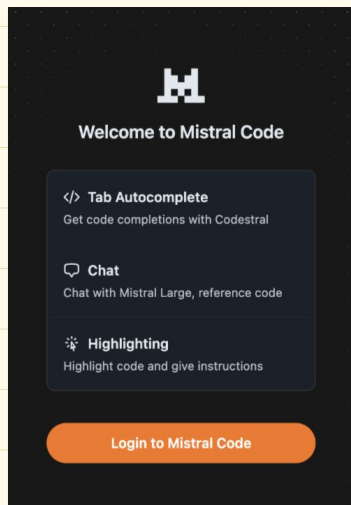
An experience shaped around your users

Memories, customizable response tone, and project creation — ensuring every user's journey is unique and personal with clear privacy settings

Memory 1

How would you introduce me if I were giving a speech?

Mistral Code Enterprise is built to solve your coding challenges



SOTA coding assistance and agentic coding, in your control.

Mistral Medium



Codestral



Devstral

Frontier coding models, fully customizable and tunable to your codebase

Generative IDE plugin

Completion

File editing

Ter

AI Foundation

Intelligence

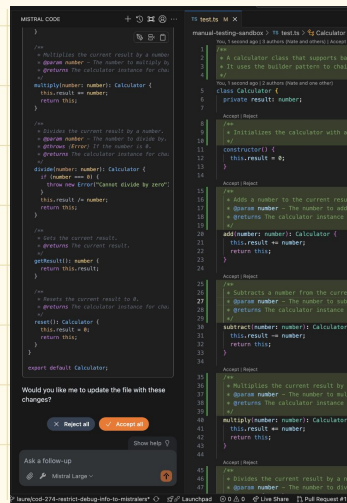
Frontier coding models

Embe

Operations

Inference infrastructure

One platform, one vendor, complete visibility across the full software stack



Deep code understanding, generation, and completion

We decided to build our own AI cluster and services

Objective: Build 100 MW compute capacity in Europe over the next 18 months

Capitalizing on both our experience with cloud providers and operating our first GPU cluster in France

Our target offering will encompass:

46,000 chips
mixing **GB200** and **GB300**

over **2 sites** 🇫🇷 / 🇪🇺

3 months between chips
delivery and capacity in
service



February 2025

Mistral Compute Overview

Bare Metal

- Pure capacity access for customers that require full control

Managed Infra

- Virtualization and orchestration in a private environment

Private AI Studio

- Developer environment with model APIs and SDKs within private clusters
- Privately deployable AI-powered products
- IT toolchain

AI Studio

- APIs, products and services for customers to experiment or get started quickly
- Familiar PaaS / SaaS toolchain

Mistral Compute Flexibility for customers to build AI on their own terms



Managed Software Services

- Managed Kubernetes
- Managed Slurm
- Bare Metal

Multiple options of training environments

Infrastructure Services

- GB200/300
- NVLink
- Infiniband
- Bluefield DPUs
- Lustre
- Vast

Advanced infrastructure and file management systems

Observability

- Stackstorm
- Grafana

Comprehensive observability stack

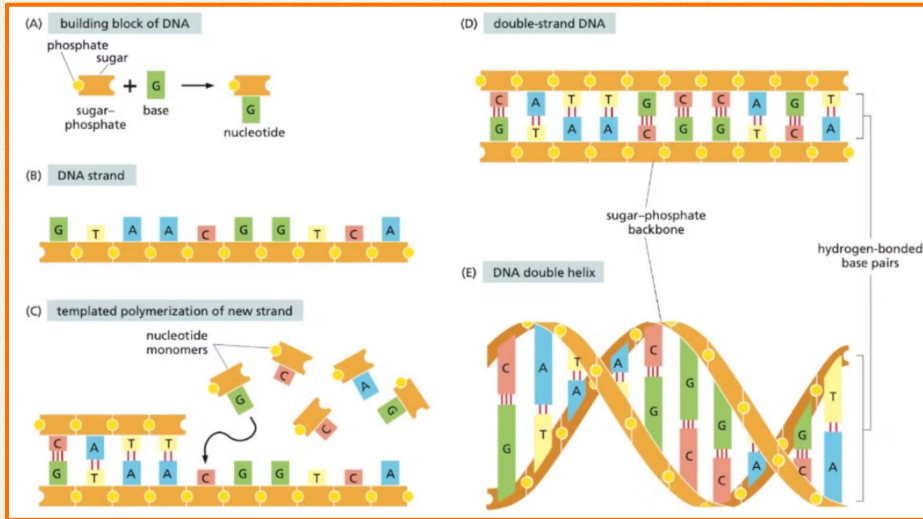


Mistral OCR & Vision

Use Cases

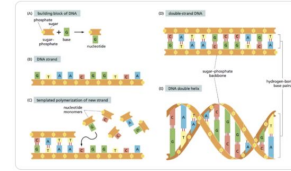
- Image captioning
- Image-to-Text Transcription (OCR)
- Data Extraction and Processing
- Analysis of complex images
- Personal Assistants (visual assistance)

Explain scientific content



GL

Explain this image



The image provides a detailed overview of the structure and replication of DNA. Here is a step-by-step summary of each part of the image:

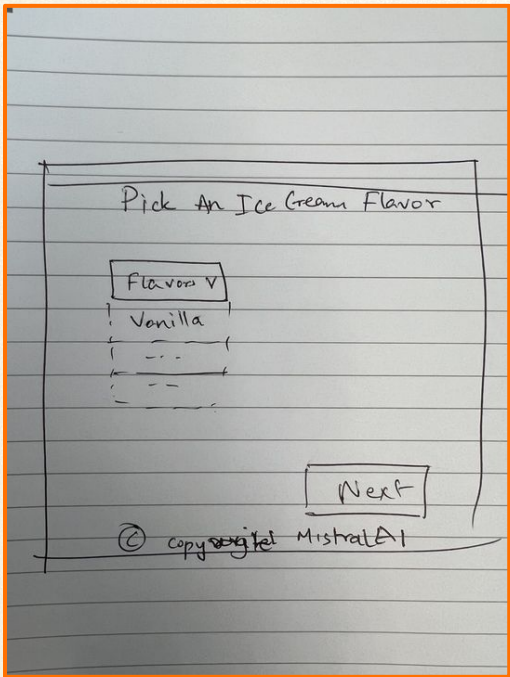
(A) Building Block of DNA:

- DNA is composed of nucleotides, which are the basic building blocks.
- Each nucleotide consists of three main components:
 - A phosphate group (yellow circles)
 - A sugar molecule (orange)
 - A nitrogenous base (colored squares: G, T, A, C)

(B) DNA Strand:

- A DNA strand is a linear sequence of nucleotides.
- The strand shown consists of the following sequence: G, T, A, A, C, G, G, T, C, A.
- The sugar-phosphate backbone is depicted as a continuous chain of yellow circles and orange sugar molecules.

Reasoning over input images



GL write HTML code to generate this webpage



Certainly! Below is an HTML code snippet that generates a webpage similar to the sketch you provided.

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Pick an Ice Cream Flavor</title>
  <style>
    body {
      font-family: Arial, sans-serif;
      display: flex;
      flex-direction: column;
      align-items: center;
      justify-content: center;
      height: 100vh;
      margin: 0;
    }
  
```

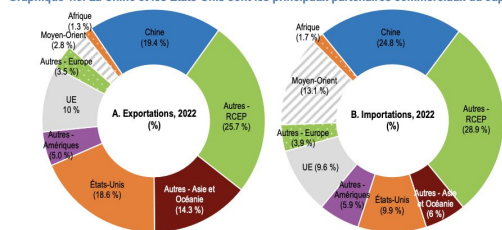
Pick An Ice Cream Flavor

- ☐ Vanilla
- ☐ Chocolate
- ☐ Strawberry

Next

© copyright MistralAI

Graphique 1.8. La Chine et les États-Unis sont les principaux partenaires commerciaux du Japon



Note : Les données concernent les échanges de biens. Les autres parties à l'Accord de partenariat économique global régional (RCEP) sont l'Australie, le Brunei Darussalam, le Cambodge, la Corée du Sud, l'Indonésie, le Laos, la Malaisie, le Myanmar, la Nouvelle-Zélande, les Philippines, Singapour, la Thaïlande et le Viet Nam. La Russie et l'Ukraine sont incluses dans l'agrégat « Autres pays d'Europe ».

Source : Ministère japonais des Finances ; et calculs de l'OCDE.

StatLink <https://statlink.b07eu>

Tableau 1.1. La croissance devrait se modérer

	2020	2021	2022	2023	2024	2025
	Prix courants (milliards de milliards JPY)	Variation annuelle, en volume (prix de 2015)				
PIB aux prix du marché	539.8	2.6	1.0	1.9	1.0	1.1
Consommation privée	291.8	0.8	2.2	0.9	1.0	0.6
Consommation publique	113.2	3.4	1.7	0.8	-1.0	0.0
Formation brute de capital fixe	137.6	-0.1	-1.4	1.8	2.6	2.8
Demande intérieure finale	542.6	1.1	1.1	1.1	1.0	1.0
Variation des stocks ¹	-1.3	0.4	0.3	0.0	-0.3	0.0
Demande intérieure totale	541.3	1.5	1.5	1.1	0.7	1.0
Exportations de biens et services	83.8	11.9	5.3	2.2	2.9	2.4
Importations de biens et services	85.3	5.1	7.9	-1.1	1.8	2.0
Solde extérieur ²	-1.5	1.0	-0.5	0.8	0.2	0.1
Pour mémoire						
PIB potentiel	-	0.7	0.6	0.4	0.4	0.4
Déficit du PIB	-	-0.2	0.3	3.8	2.7	2.2
Écart de production	-	-1.9	-1.5	-0.1	0.5	1.2
Indice des prix à la consommation (IPC)	-	-0.2	2.5	3.2	2.6	2.0
IPC sous-jacent ²	-	-0.7	0.3	2.7	2.3	2.0
Taux de chômage (% de la population active)	-	2.8	2.6	2.6	2.5	2.4
Taux d'épargne des ménages, net (% du revenu disponible)	-	7.8	5.2	2.8	2.3	0.7
Solde financier des administrations publiques (% du PIB)	-	-6.2	-5.7	-5.2	-4.3	-3.3
Dettes brutes des administrations publiques (% du PIB)	-	239.3	244.8	243.5	243.3	242.3
Solde des paiements courants (% du PIB)	-	3.9	1.8	3.4	3.8	3.9

1. Contributions aux variations du PIB réel ; montant effectif dans la première colonne.

2. Indice des prix à la consommation hors produits alimentaires et énergie.

Source : OCDE, Perspectives économiques de l'OCDE, n° 114 (base de données), données actualisées après la publication des comptes nationaux du Japon le 8 décembre 2023.

ÉTUDES ÉCONOMIQUES DE L'OCDE : JAPON 2024 (VERSION ABRÉGÉE) © OCDE 2024

Quelles sont les grandes tendances économiques pour le Japon?

Les grandes tendances économiques pour le Japon peuvent être analysées à partir des données fournies et des contextes économiques globaux. Voici quelques points clés :

1. ****Partenaires Commerciaux****

- ****Chine et États-Unis**** : Les États-Unis et la Chine sont les principaux partenaires commerciaux du Japon, ce qui signifie que les fluctuations économiques dans ces deux pays peuvent avoir un impact significatif sur l'économie japonaise.

- ****RCEP**** : Le Japon est également impliqué dans l'Accord de partenariat économique global régional (RCEP), qui inclut plusieurs pays d'Asie-Pacifique. Cela ouvre des opportunités pour diversifier les échanges commerciaux et réduire la dépendance vis-à-vis de quelques partenaires clés.

2. ****Croissance Économique****

- ****Modération de la Croissance**** : La croissance économique du Japon devrait se modérer au cours des prochaines années. Les prévisions montrent une croissance annuelle en volume de 1.0% à 1.1% entre 2023 et 2025.

- ****PIB et Consommation**** : La consommation privée et publique, ainsi que la formation brute de capital fixe, montrent des variations modérées, ce qui indique une croissance économique stable mais non explosive.

3. ****Inflation et Taux de Changement****

- ****Inflation**** : L'indice des prix à la consommation (IPC) montre une augmentation modérée, passant de 2.5% en 2022 à 2.0% en 2025. L'IPC sous-jacent, qui exclut les produits alimentaires et l'énergie, suit une tendance similaire.

- ****Taux de Changement**** : Le taux de chômage reste stable autour de 2.5% à 2.6%, ce qui indique un marché du travail relativement stable.

4. ****Épargne et Dette Publique****

- ****Épargne des Ménages**** : Le taux d'épargne des ménages diminue progressivement, passant de 7.8% en 2020 à 0.7% en 2025.

- ****Dette Publique**** : La dette brute des administrations publiques reste élevée, autour de 240% du PIB, mais montre une légère diminution au fil des années.

5. ****Solde des Paiements Courants****

- ****Solde Positif**** : Le solde des paiements courants reste positif, autour de 3.4% à 3.9% du PIB, ce qui indique une économie relativement stable en termes de flux financiers internationaux.

6. ****Défis et Opportunités****

- ****Défis**** : Le Japon fait face à des défis démographiques, notamment un vieillissement de la population et une baisse de la population active, ce qui pourrait affecter la croissance économique à long terme.

- ****Opportunités**** : La participation à des accords commerciaux comme le RCEP offre des opportunités pour diversifier les échanges commerciaux et renforcer les relations économiques avec d'autres pays de la région.

Conclusion

Le Japon continue de naviguer dans un environnement économique global complexe, avec des partenaires commerciaux clés et des accords régionaux qui offrent à la fois des défis et des opportunités. La croissance économique est modérée, avec une inflation contrôlée et un marché du travail stable. Cependant, des défis démographiques et une dette publique élevée restent des préoccupations majeures pour l'avenir économique du Japon.

Performed with OCR API + Mistral Small 3

"[New Beginnings] is the anti-prison," Vincent N. Schiraldo, director of the Department of Youth Rehabilitation Services, told *The Washington Post* in May. "What we had before was a training school for them to become adult inmates. We want them to aspire to college, to be in a place that looks like you care about them."

In August, California state officials announced plans to close Herman G. Stark Youth Correctional Facility in China. A report released in 2007 concluded that the environment was so bad at the facility that youth were especially prone to violence or suicide.

The Community Justice Network for Youth (CJNY), a juvenile justice advocacy group based in California, warned that the Stark closure is bittersweet, adding that the state still intends to transform the facility into an adult prison.

RICHMOND, CALIF. STUDENTS LEARN FROM TRAGEDY

The horrific news of a 16-year-old girl who was allegedly gang-raped outside of her homecoming dance at Richmond High School in Northern California shocked millions. What made it even worse were reports that the attack was witnessed by over a dozen people who, over the course of two hours, allegedly took photos and joined the attack, but failed to intervene or call police.

In the weeks that followed, Richmond High students faced intense media scrutiny. The students, most of whom are working class and of color, were called "animals" and "monsters" by several outraged media observers.

Students acted quickly. With the help of campus-based organizations like Youth Together, a Bay Area education reform organization, hundreds of dedicated students and teachers mobilized candlelight vigils and financial support funds to help the victim recover.

Students and activists are also developing gender violence trainings to be added into the school's permanent curriculum. By the end of the training, organizers hope that students will examine how they perpetuate violence in their own lives, know how to respond to a bystander and become certified anti-violence trainers.

WISCONSIN STUDENTS DREAM BIG

While federal officials stalled on immigration reform this year, students in Wisconsin went full steam ahead when they successfully passed a state-based version of the DREAM Act.

According to immigration activists, each year thousands of undocumented students are barred from going to college because they don't qualify for state or federal financial aid. Student-led groups across the country have increased their efforts to pass the DREAM Act, a proposed piece of federal legislation that would provide undocumented students with a path toward legalization and qualify them for financial aid.

On June 29, Wisconsin became the 11th state in the nation to allow undocumented students to pay in-state tuition. Students organized with the support of immigrant advocacy groups Voces de la Frontera and Students United for Recognizing Immigrant Rights (SUFIR).

"[New Beginnings] is the anti-prison," Vincent N. Schiraldo, director of the Department of Youth Rehabilitation Services, told *The Washington Post* in May. "What we had before was a training school for them to become adult inmates. We want them to aspire to college, to be in a place that looks like you care about them."

In August, California state officials announced plans to close Herman G. Stark Youth Correctional Facility in China. A report released in 2007 concluded that the environment was so bad at the facility that youth were especially prone to violence or suicide.

The Community Justice Network for Youth (CJNY), a juvenile justice advocacy group based in California, warned that the Stark closure is bittersweet, adding that the state still intends to transform the facility into an adult prison.

RICHMOND, CALIF. STUDENTS LEARN FROM TRAGEDY

The horrific news of a 16-year-old girl who was allegedly gang-raped outside of her homecoming dance at Richmond High School in Northern California shocked millions. What made it even worse were reports that the attack was witnessed by over a dozen people who, over the course of two hours, allegedly took photos and joined the attack, but failed to intervene or call police.

In the weeks that followed, Richmond High students faced intense media scrutiny. The students, most of whom are working class and of color, were called "animals" and "monsters" by several outraged media observers.

Students acted quickly. With the help of campus-based organizations like Youth Together, a Bay Area education reform organization, hundreds of dedicated students and teachers mobilized candlelight vigils and financial support funds to help the victim recover.

Students and activists are also developing gender violence trainings to be added into the school's permanent curriculum. By the end of the training, organizers hope that students will examine how they perpetuate violence in their own lives, know how to respond with a bystander and become certified anti-violence trainers.

WISCONSIN STUDENTS DREAM BIG

While federal officials stalled on immigration reform this year, students in Wisconsin went full steam ahead when they successfully passed a state-based version of the DREAM Act.

According to immigration activists, each year thousands of undocumented students are barred from going to college because they don't qualify for state or federal training. 100 Student-led groups across the country have increased their efforts to pass the DREAM Act, a proposed piece of federal legislation that would provide undocumented students with a path toward legalization and qualify them for financial aid.

On June 29, Wisconsin became the 11th state in the nation to allow undocumented students to pay in-state tuition. Students organized with the support of immigration advocacy groups Voces de la Frontera and Students United for Recognizing Immigrants' Rights (NUPATR).

Mistral Workflows: : Orchestrate AI Workflows with confidence

Accelerate. Observe. Trust. Scale.

Accelerate AI Use Cases



Provides a Python-first framework with pre-built components for rapid development

Launch multi-step AI processes (e.g., document processing, knowledge extraction) in days, not months.

End-to-End Observability



Built-in OpenTelemetry integration and real-time monitoring dashboards

Track workflows with Mistral Workflows UI - see every step, retry, and failure in one place

Reliable by Design



Automatic retries, state persistence, and fault tolerance powered by Temporal

Workflows resume exactly where they left off after a crash

Effortless Scalability



Mistral Workflows decouples logic from infrastructure, enabling horizontal scaling without refactoring.

Handle 10x workload by adding workers, not changing code

Automating Customer Care for a Major Transport Company

Our client's critical pain point

Highly inefficient RFI process involving 10 people during 2 weeks



How we solved it with AI?

RAG-powered agent for full RFI automation

- › **Unified interface** allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process
- › **Improved quality and accuracy** of answers thanks to a unified RFI database

Key Outcomes

E2E automation of 1,000 RFI yearly

Provide sharper and more consistent answers

Allow higher conversion rate and win rate

Automating Customer Care for a Major Transport Company

We identify our client's pain point

FROM Highly inefficient RFI process involving 10 people during 2 weeks

- > **Repetitive and inefficient tasks** leading to duplicated efforts and inconsistent outcomes
- > **Fragmented and inaccessible data**, slowing down decision-making and limiting visibility

We deploy our transforming solution

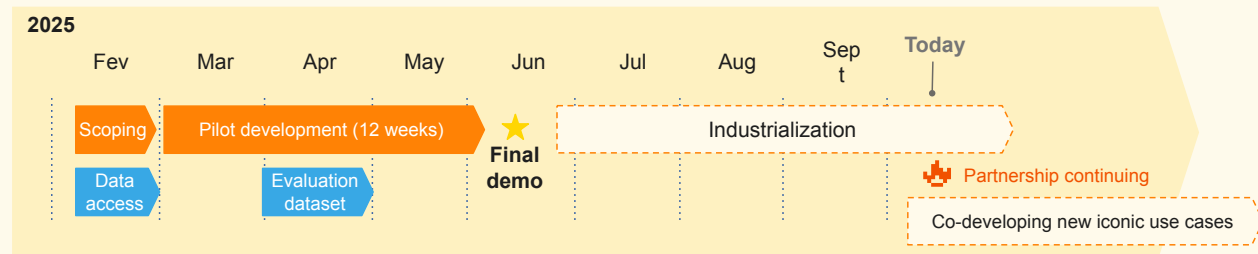
TO RAG-powered agent for full RFI automation

- > **Unified interface** allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process
- > **Improved quality and accuracy** of answers thanks to a unified RFI database

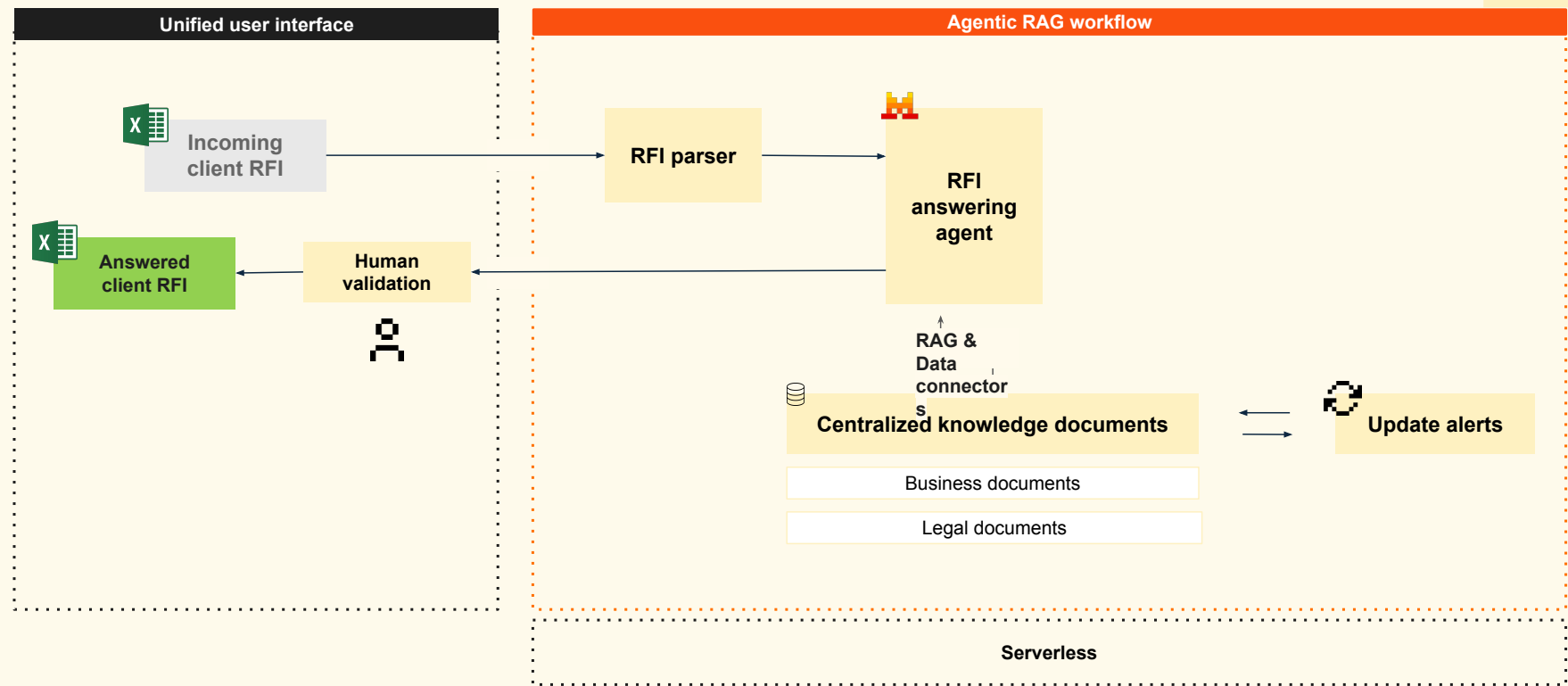
■ Key outputs

- > E2E automation of **1,000 RFI yearly**
- > Provide sharper and more consistent answers
- > Allow higher conversion rate and win rate

From the iconic use case to scaling the partnership together



Automating Customer Care for a Major Transport Company



Driving AI Transformation in manufacturing

Internal enablement



AI-powered access for operators to technical knowledge and critical data

Drive efficiency gains



Core Process Automation



AI for SML's product portfolio and research, development and operations

Accelerate time to market



Customer & market innovation



AI powered in-car assistant for navigation, manual Q&A and commands

Transform user's experience



Driving AI Transformation for corporate functions

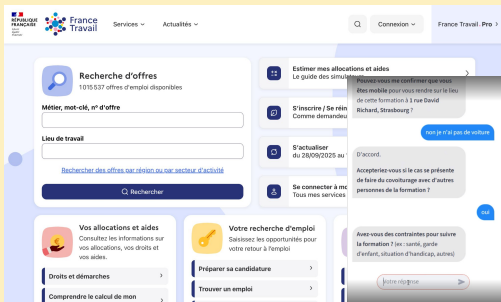
Internal enablement

Support chatbot



RAG chat agent for advisors to query France Travail's information system

Drive efficiency gains for 40k users



Empower your employees

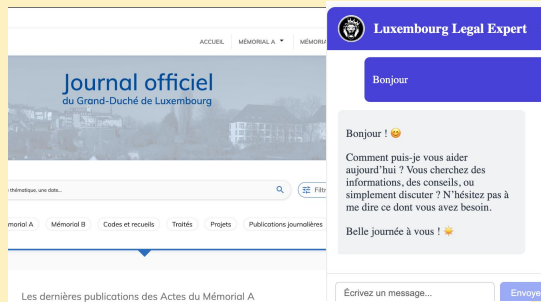
Legal Assistant



LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG

RAG legal agent on Legilux website for legal answers for experts and citizens

Enable legal data access for 600k users

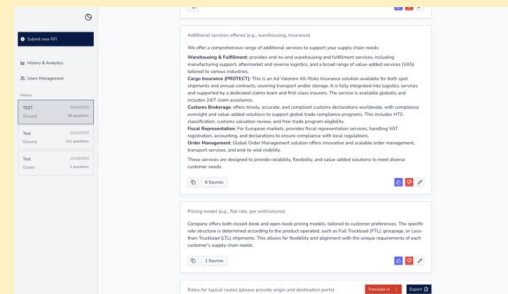


Core processes automation

Accelerate RFI process

Unified interface allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process

E2E automation of 1,000 RFI yearly



Mapping job-seekers to a standardized skills profile



We identify our client's pain point

FROM Incomplete resume using different standard not reflecting real talent skills

- **Multiple skills descriptions** and frameworks leading to difficulties to compare candidates
- **Candidates' self reflection** on their skills usually incomplete or imprecise



We deploy our transforming solution

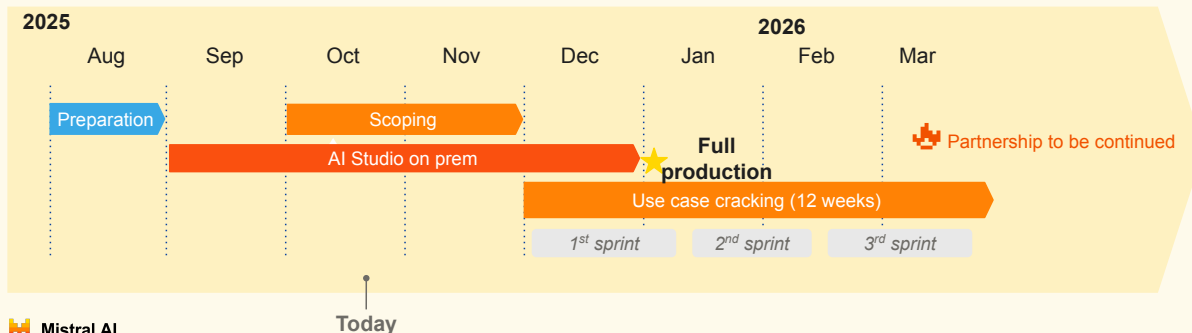
TO Unified augmented skills profile reflecting candidates true capabilities

- **Single ESCO framework** used for all countries involved (Luxembourg, France, Belgium, Germany)
- **Self-standing solution** that could be reused for multiple services

Key outputs

- Improve resume generation for more than Xk users
- Improved matching of trainings and job offers with candidates

From the iconic use case to scaling the partnership together



Automating Customer Care for a Major Transport Company

Our client's critical pain point

Highly inefficient RFI process involving 10 people during 2 weeks



How we solved it with AI?

RAG-powered agent for full RFI automation

- › **Unified interface** allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process
- › **Improved quality and accuracy** of answers thanks to a unified RFI database

Key Outcomes

E2E automation of 1,000 RFI yearly

Provide sharper and more consistent answers

Allow higher conversion rate and win rate

Automating Customer Care for a Major Transport Company

We identify our client's pain point

FROM Highly inefficient RFI process involving 10 people during 2 weeks

- > **Repetitive and inefficient tasks** leading to duplicated efforts and inconsistent outcomes
- > **Fragmented and inaccessible data**, slowing down decision-making and limiting visibility

We deploy our transforming solution

TO RAG-powered agent for full RFI automation

- > **Unified interface** allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process
- > **Improved quality and accuracy** of answers thanks to a unified RFI database

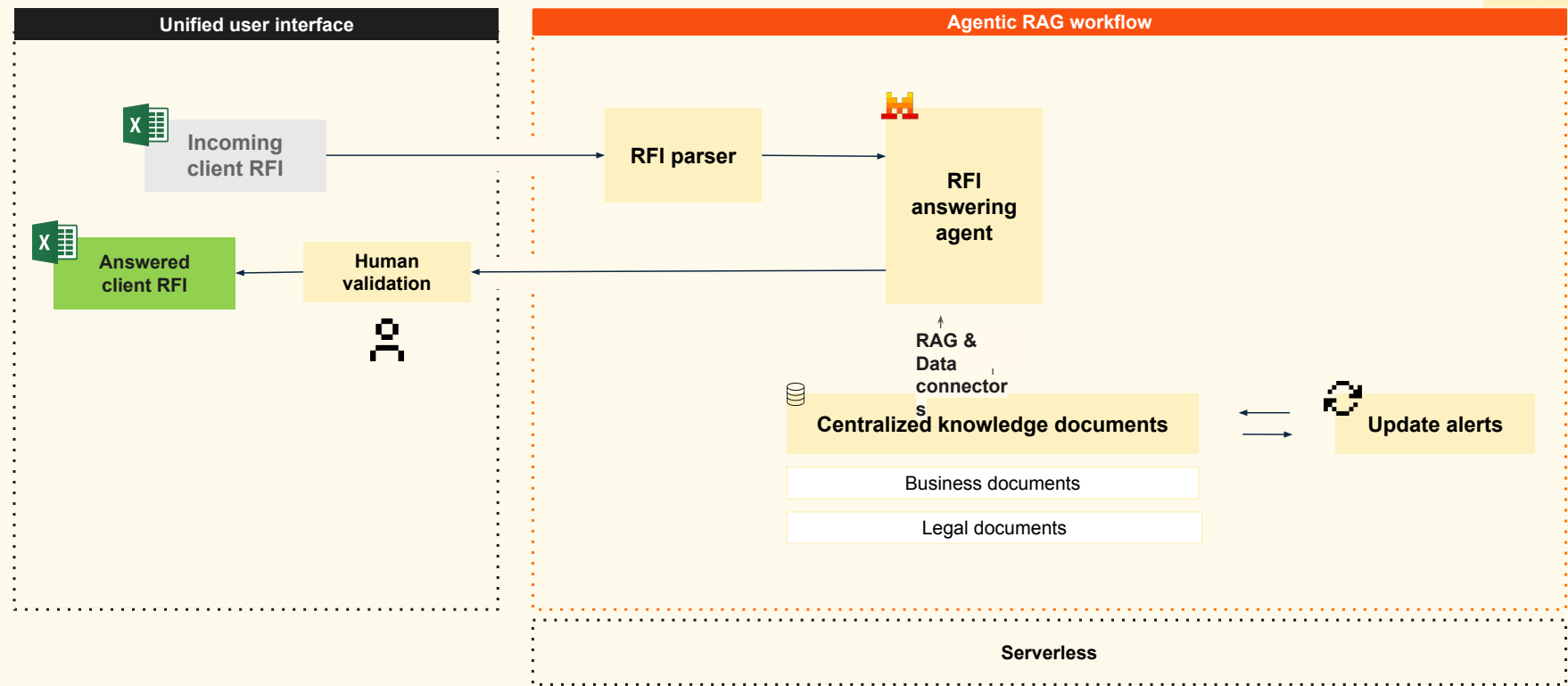
■ Key outputs

- > E2E automation of **1,000 RFI yearly**
- > Provide sharper and more consistent answers
- > Allow higher conversion rate and win rate

From the iconic use case to scaling the partnership together



Automating Customer Care for a Major Transport Company



Driving AI Transformation for defense and aerospace industry

Robotics & autonomy



Advanced research themes such as multimodal models, robotics, automation, and embedded systems.

Industrialization of key products

Custom security LLM



End-to-end process of building custom models to enhance security tasks capabilities

Power operational agility

Custom VLAs



Customized Vision-Language-Action model that takes multiple inputs (video, text) and outputs machine actions

Enhance decision-making on the field

Effective AI Agent Customer Support for Major Retail Bank

Our client's critical pain point

Manual and time-consuming banking processes on the app or through advisors



How we solved it with AI?

In-app AI Banking Chatbot for customer support to fully automate user journey

- › **Single point of contact via chatbot** to automate customer requests with fast and accurate responses
- › **Integration with loyalty accounts** allows customers to view their account details and manage them directly

Key Outcomes

E2E automation of customer support requests for **1M+ active users** on their banking app

Improved **user experience**

Effective AI Agent Customer Support for Major Retail Bank

We identify our client's pain point

FROM Manual and time-consuming banking processes on the app or through advisors

- > Massive volumes of repetitive requests are received daily by customer support teams
- > Unclear steps and failed attempts degrade customer experiences

We deploy our transforming solution

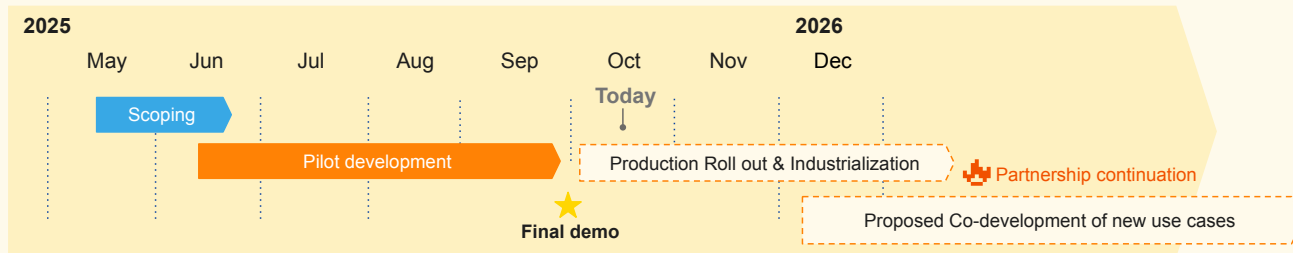
TO In-app AI Banking Chatbot for customer support to fully automate user journey

- > Full end-to-end automation of user journeys without advisor intervention on the app
- > Data retrieval pipeline enabling the chatbot to perform actions on behalf of users (e.g., search payments, unblock cards, etc.)

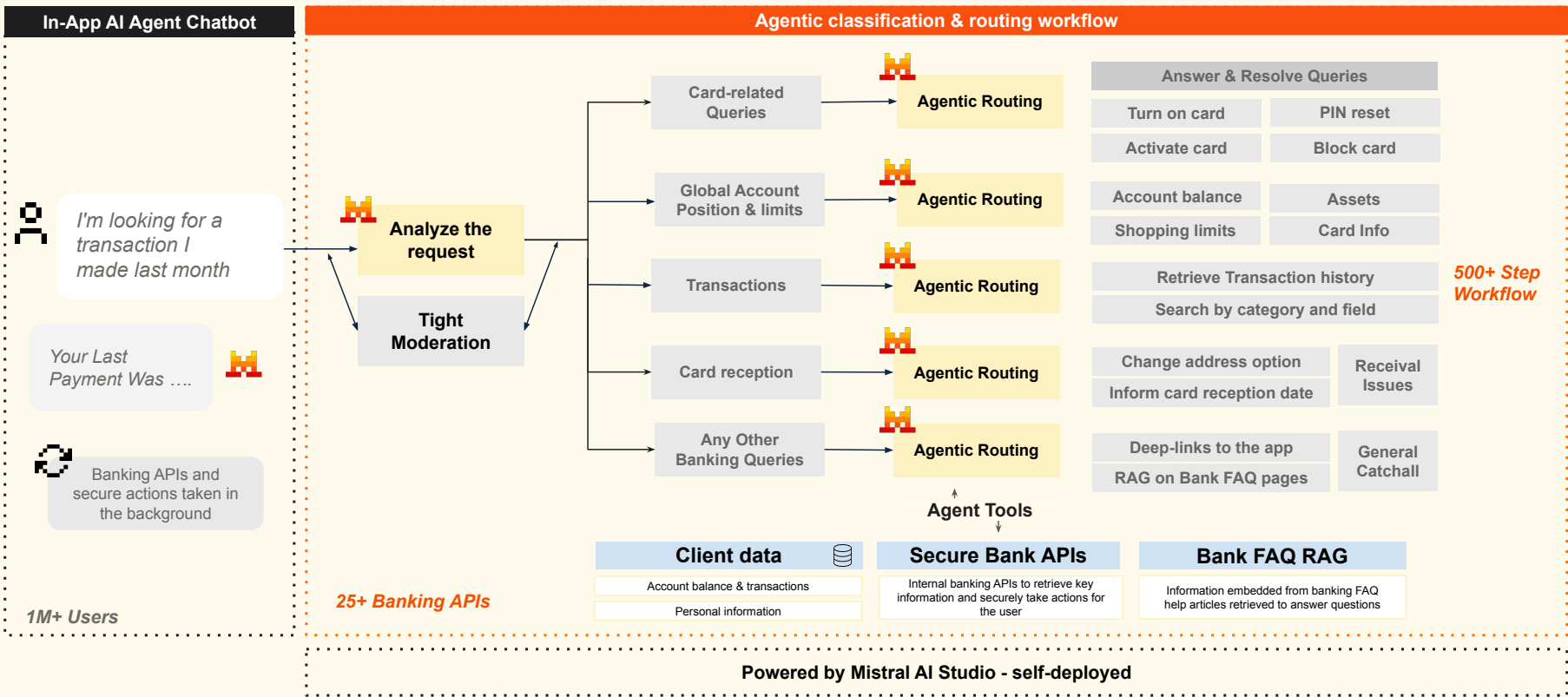
■ Key outputs

- > E2E automation of **customer support requests for 1M+ active users on their banking app**
- > Improved user experience
- > Going to production at scale with their user-base

From the iconic use case to scaling the partnership together



Effective AI Agent Customer Support for Major Retail Bank



Transforming Customer Support at a Major Retail Bank

Our client's critical pain point

Fragmented and inefficient digital client support creating client frustration



How we solved it with AI?

Augmented chat agent allowing a seamless experience on the app

- › Agentic workflows **automates requests** through a RAG pipeline that ensures accurate, cited responses
- › Complex requests requiring human intervention are qualified and **routed toward an advisor**

Automation of 200k + requests per month from 160 different requests categories

Reduced request processing time

Enhanced customer satisfaction

Transforming Customer Support at a Major Retail Bank

We identify our client's pain point

FROM Fragmented and inefficient digital client support creating client frustration

- Massive volumes of repetitive requests are received through shattered points of contact
- Delays in resolving requests lead to frustration and a **poor customer experience**

We deploy our transforming solution

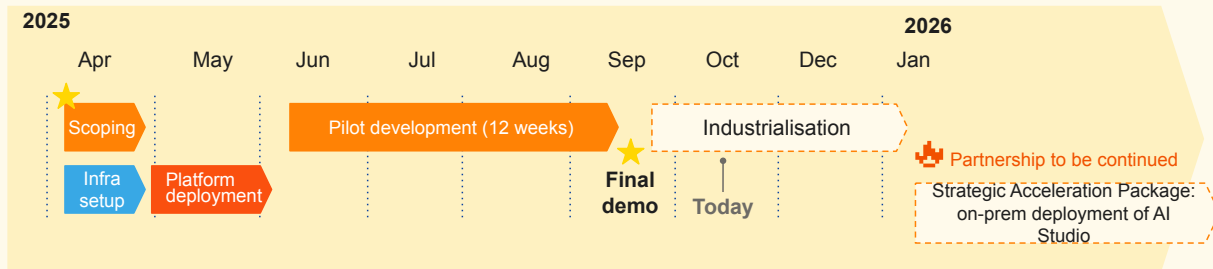
TO Augmented chat agent allowing a seamless experience on the app

- Agentic workflows **automates requests** through a RAG pipeline that ensures accurate, cited responses
- Complex requests requiring human intervention are qualified and **routed toward an advisor**

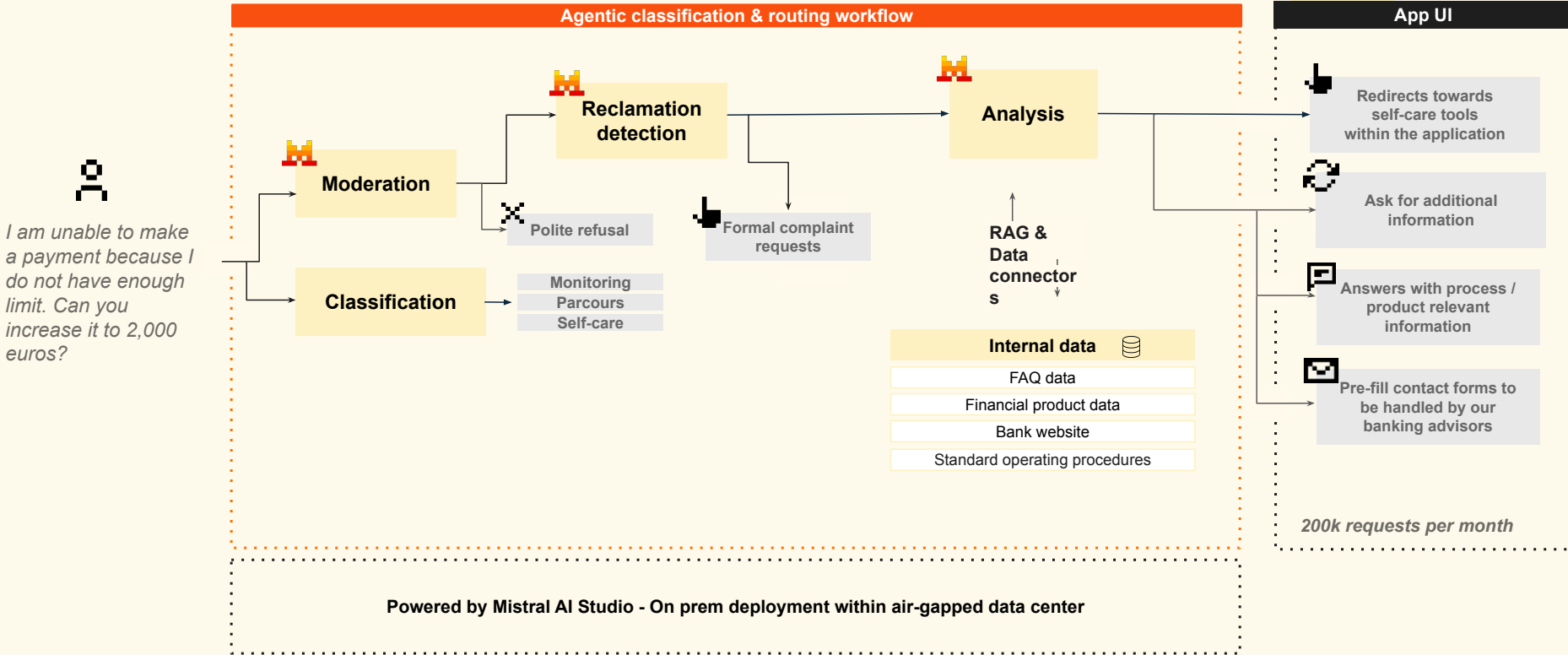
■ Key outputs

- Handle **200k + requests per month** from 160 different requests categories
- Reduce request processing time
- Enhance customer satisfaction

From the iconic use case to scaling the partnership together



Transforming Customer Support at a Major Retail Bank



Enabling Fact Checking for a Large Media Company

Our client's critical pain point

Tedious manual fact checking process with high risk of mistakes and hallucinations

How we solved it with AI? Fact-checking Productivity Agent

- › **Unified interface** allowing subject matter experts to fill RFI in faster and more efficient way thanks to a standardized process
- › Adapting to institutional policies and **ensuring the credibility of publications through** carefully selected data sources

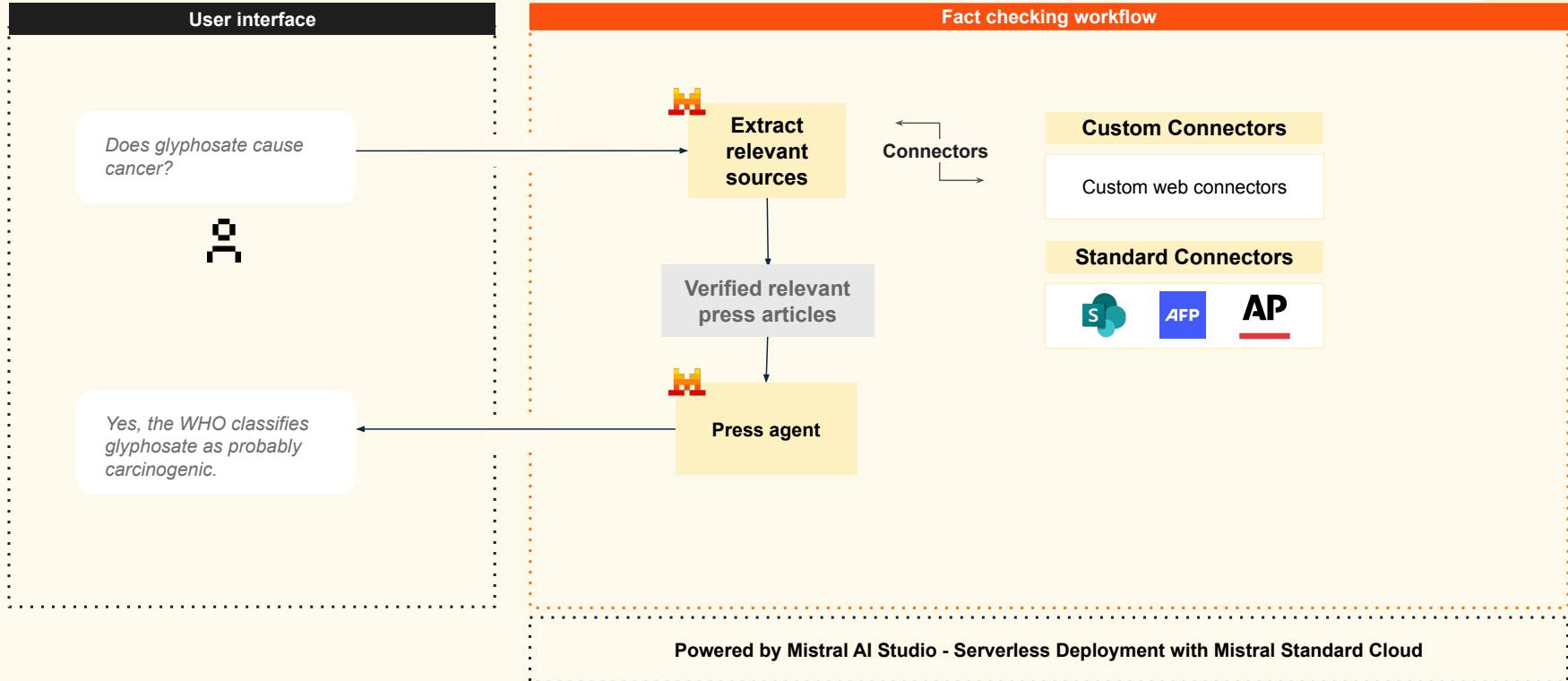
Key Outcomes

Faster, more efficient fact-checking in journalism

Reduced time spent on manual verification

Increased confidence in the accuracy of published information

Enabling Fact Checking for a Large Media Company



Democratizing Legal Knowledge for All Citizens

Our client's critical pain point **Fragmented access to legal knowledge** for citizens and experts

 **How we solved it with AI?** **RAG augmented legal expert agent** fully integrated into government website

- › Single point of access for experts and citizen to provide **reliable and up to date legal information**
- › **Co-built with legal experts** to ensure model relevance and safety

Key Outcomes

Legal knowledge access for 600k + yearly users on the website

Improved administrative processes for the private sector and citizens

Democratizing Legal Knowledge for All Citizens

We identify our client's pain point

FROM Fragmented access to legal knowledge for citizens and experts

- Limited access to legal knowledge leads to **frustration and confusion** among citizens
- **Increased burden on public administrations** to manage citizens' legal uncertainties and missteps

We deploy our transforming solution

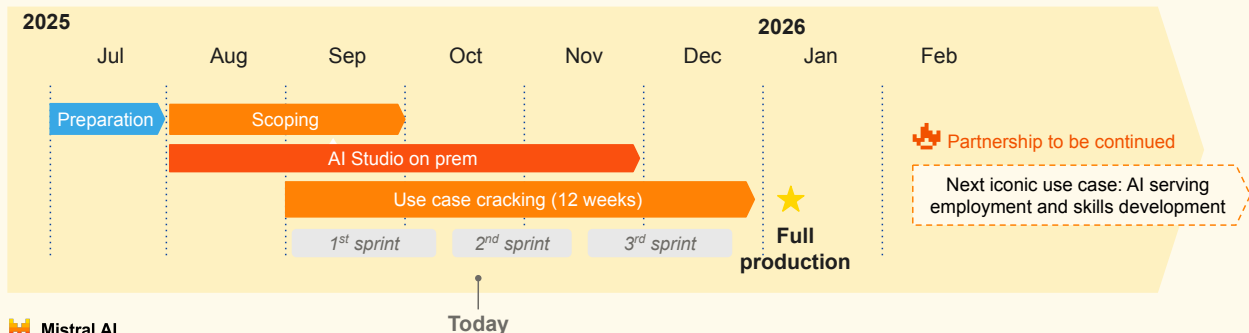
TO RAG augmented legal expert agent fully integrated into government website

- Single point of access for experts and citizen to provide **reliable and up to date legal information**
- **Co-built with legal experts** to ensure model relevance and safety

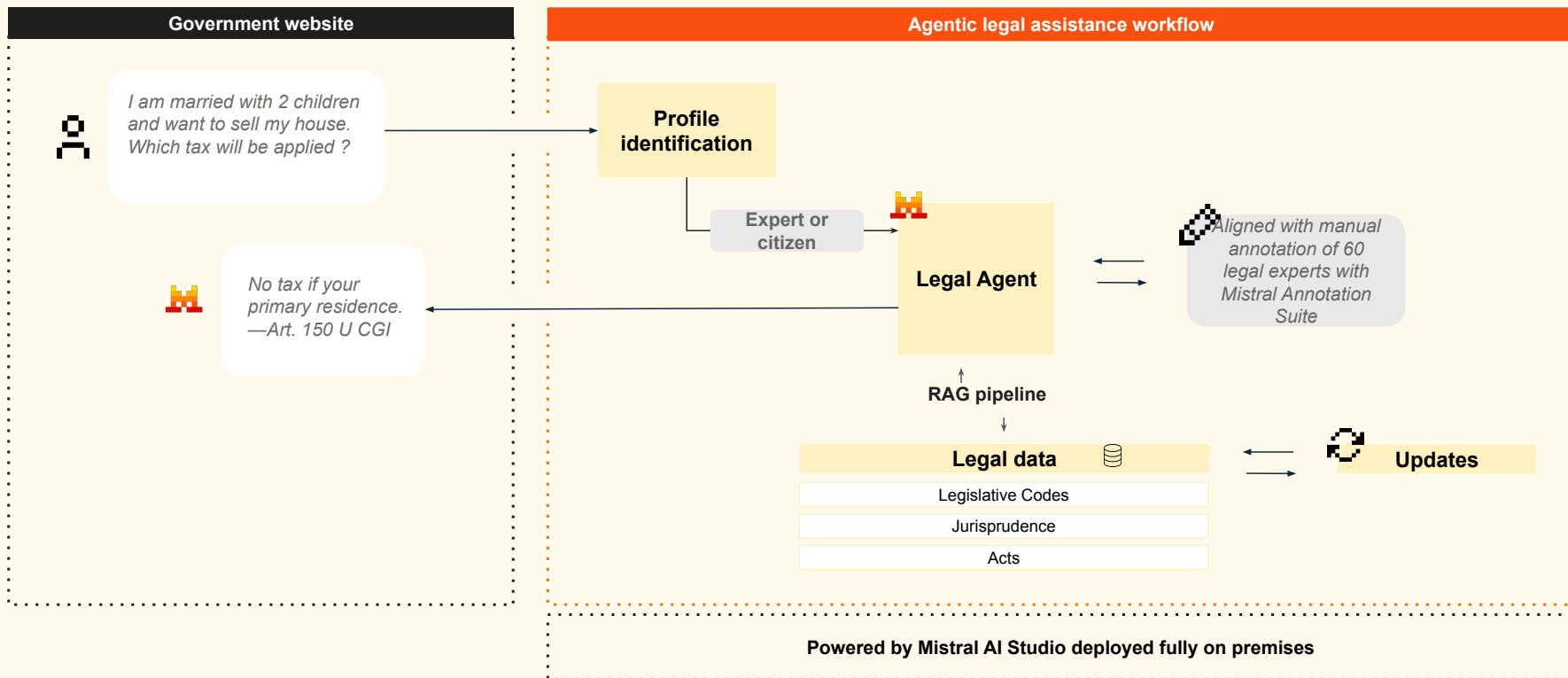
■ Key outputs

- Legal knowledge access for **600k + yearly users** on the website
- Improved administrative processes for the private sector and citizens

From the iconic use case to scaling the partnership together

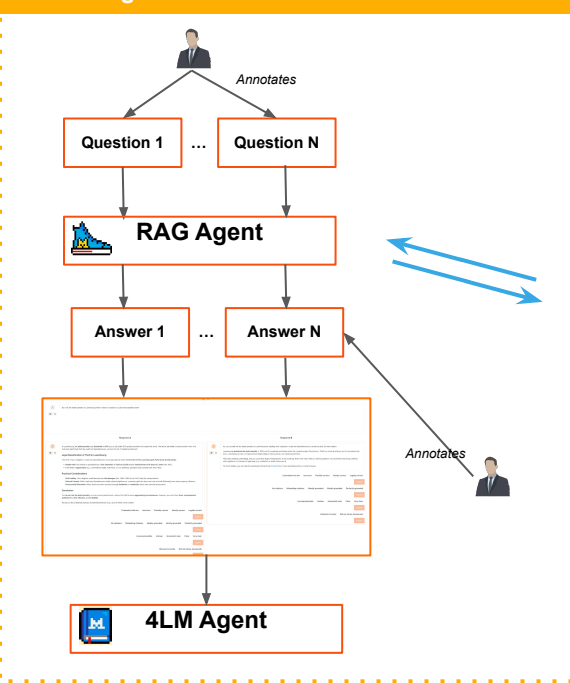


Democratizing Legal Knowledge for All Citizens

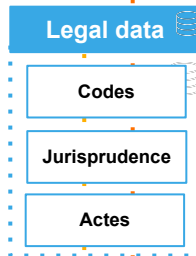
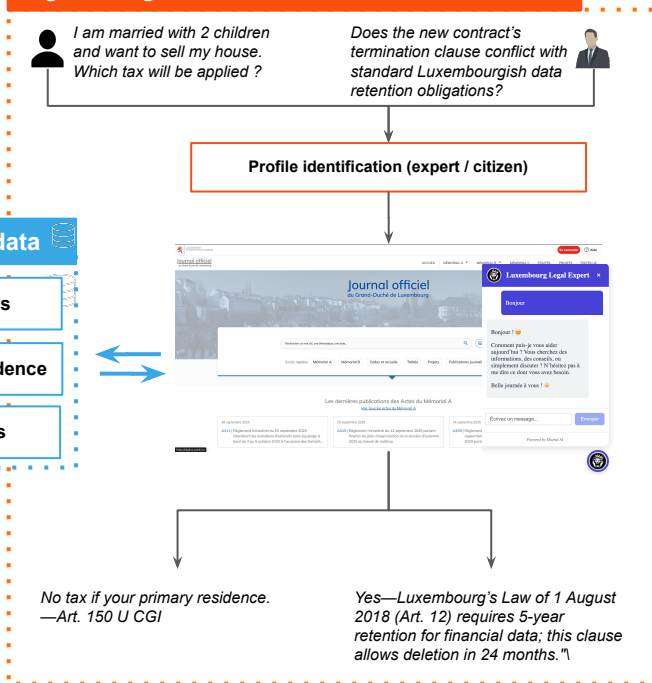


Luxembourg Legal LLM (4LM): Making legal knowledge accessible for citizens

Model alignment workflow



Agentic legal assistance workflow



Key takeaways

Helping up to **600K citizens** in their daily legal needs

State of the art RAG performance **mixing semantic and hierarchical search**

Fully integrated into Legilux website and grounded into live legal data

Co-built with **legal experts** to ensure model relevance and safety using our annotation suite in Mistral AI Studio

Best in class AI model for South-East-Asian languages

Context

Create SOTA model for Singaporean citizens

Why:

Some languages are under-represented in general model datasets. This forces local citizen to use english to get best performance out of AI.

Solution:

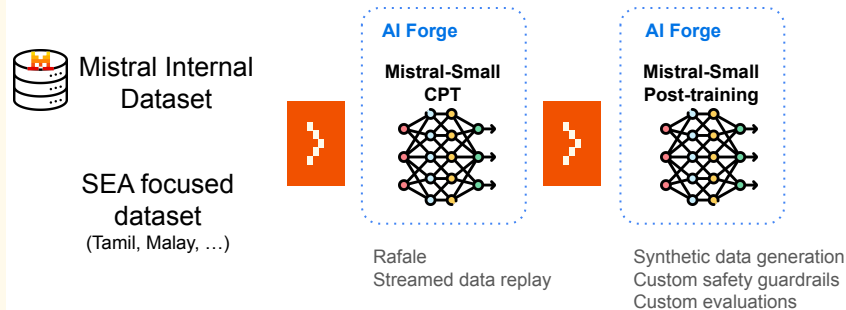
Create language focused LLMs for local languages with SOTA capabilities.

Outcome:

Public services, local companies and startups, research lab can leverage to its full power in their native languages without compromising on performance.

Co-construction with client teams

Co-training program



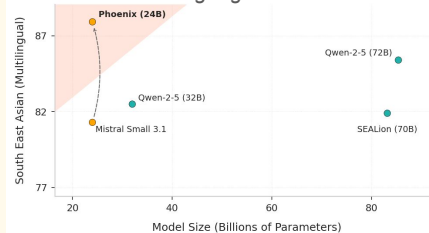
Support from Mistral Science



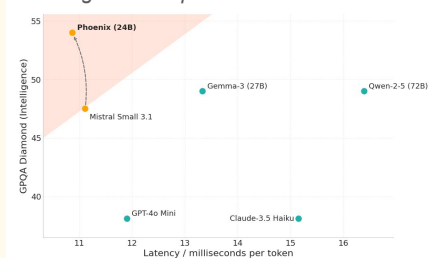
Outcomes

Phoenix

SOTA in SEA Languages



With general capabilities



Best in class AI model for Arabic language

Context

Create SOTA model for arabic language

Why:

Some languages are under-represented in general model datasets. This forces local citizen to use english to get best performance out of AI.

Solution:

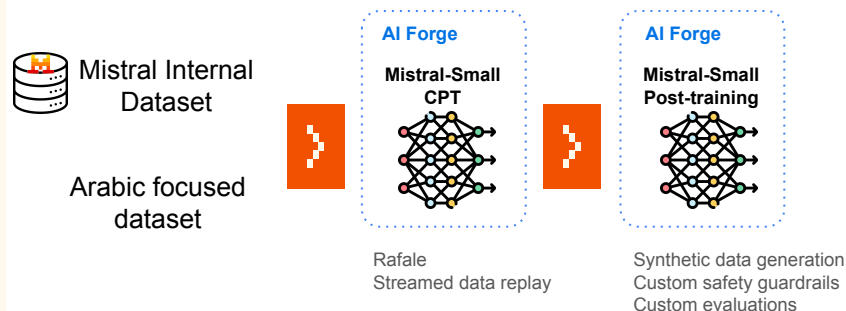
Create language focused LLMs for local languages with SOTA capabilities.

Outcome:

Public services, local companies and startups, research lab can leverage to its full power in their native languages without compromising on performance.

Co-construction with client teams

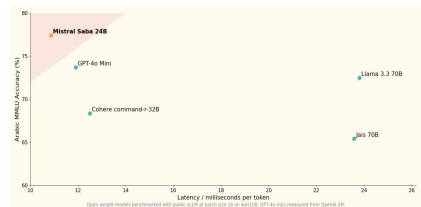
Co-training program



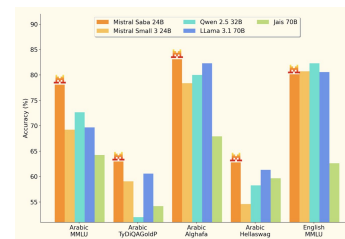
Outcomes

Saba

SOTA in arabic



With general capabilities



Powering Proactive Plant Maintenance for Veolia

Our client's critical pain point **Inefficient and costly maintenance** for water treatment systems

 **How we solved it with AI?** **AI maintenance assistant** acting as a virtual field service technician

- › **Daily systems check** to validate or recommend adjustments to control set points
- › Enabled **proactive maintenance** between scheduled visits to prevent system upsets and extend asset lifespan

Key Outcomes

+20% targeted productivity increase

Targeted **2x quicker time to resolution**

Proactive monitoring for assets longevity

Powering Proactive Plant Maintenance for Veolia

We identify our client's pain point

FROM Inefficient and costly maintenance for water treatment systems

- Troubleshooting is a tedious process consuming roughly **33% of an operation manager's time**
- **Data is scattered**, making it hard to spot issues early or act before problems arise

We deploy our transforming solution

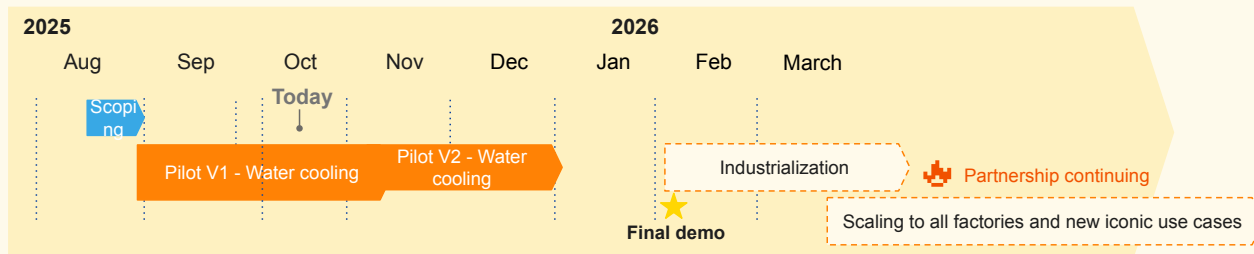
TO AI chatbot advisor acting as a virtual field service technician

- **Daily systems check** to validate or recommend adjustments to control set points
- Enabled **proactive maintenance** between scheduled visits to prevent system upsets and extend asset lifespan

■ Key outputs

- **+20% targeted productivity** increase
- Targeted **2x quicker time to resolution**
- Proactive monitoring for assets longevity

From the iconic use case to scaling the partnership together



Powering Proactive Plant Maintenance for Veolia

