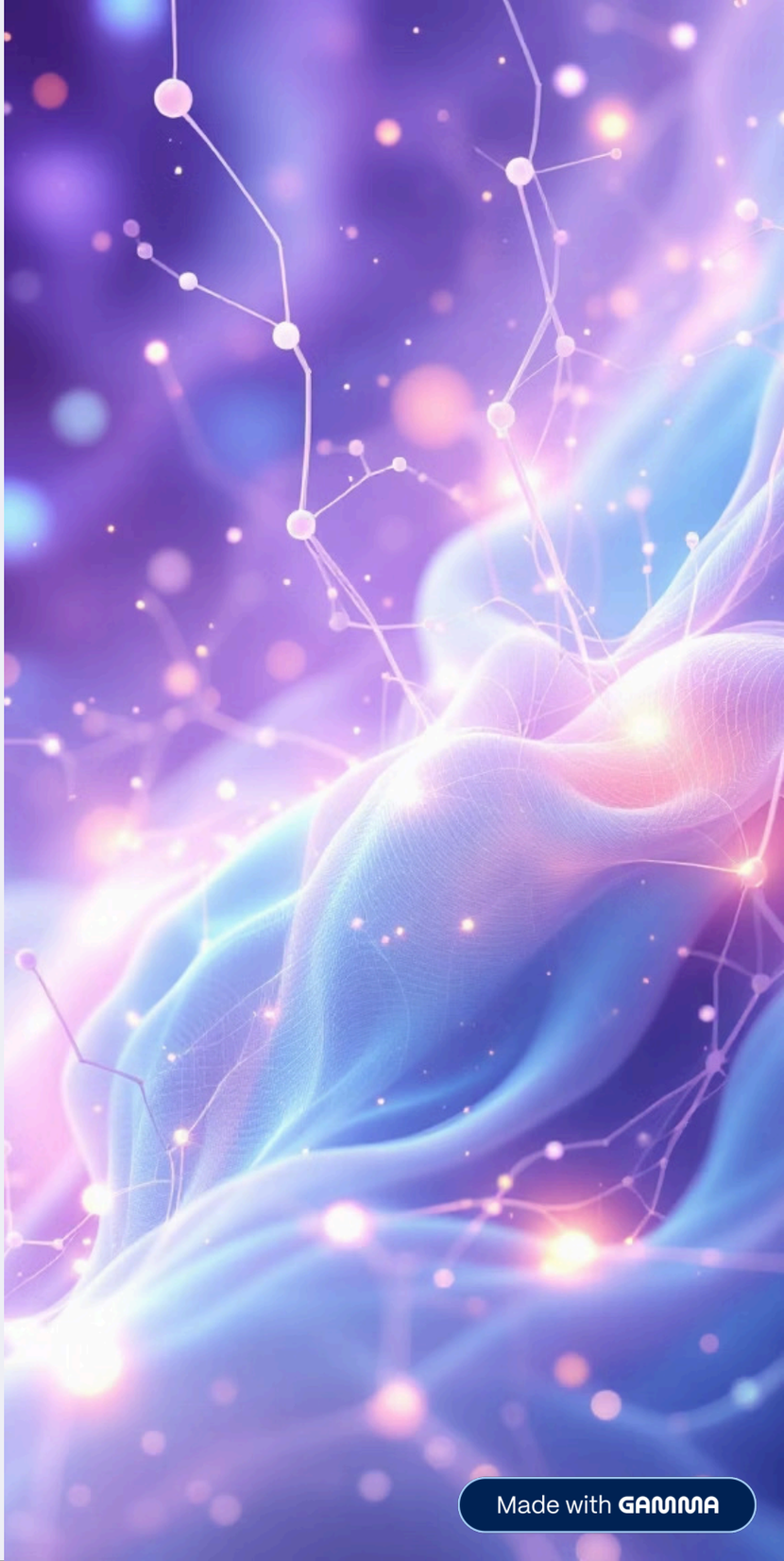
# Significance of Generative AI & Language Models

Generative AI and LLMs transform society rapidly

New opportunities and challenges emerge

Tracking the latest research is essential



# Research Summary Scope & Sources

### Research Origins

Leading industrial labs and universities from 2024 onward

### Information Sources

Only publicly available data like papers, blogs, announcements

### Report Goal

Comprehensive overview of latest generative AI developments

# Paradigm Shift & Ecosystem Dynamics

## Fast-Evolving Field

- Knowledge becomes rapidly outdated
- Ongoing vibrant research activity

## Industrial & Academic Roles

- Industry: Resources and practical innovation
- Universities: Fundamental theoretical research

# AI Model Evolution: 2025 Landscape

Key developments in AI models from OpenAI, Google DeepMind, and Anthropic.


# GPT-4.1 Series

## New Models

GPT-4.1, mini, nano versions outperform GPT-40

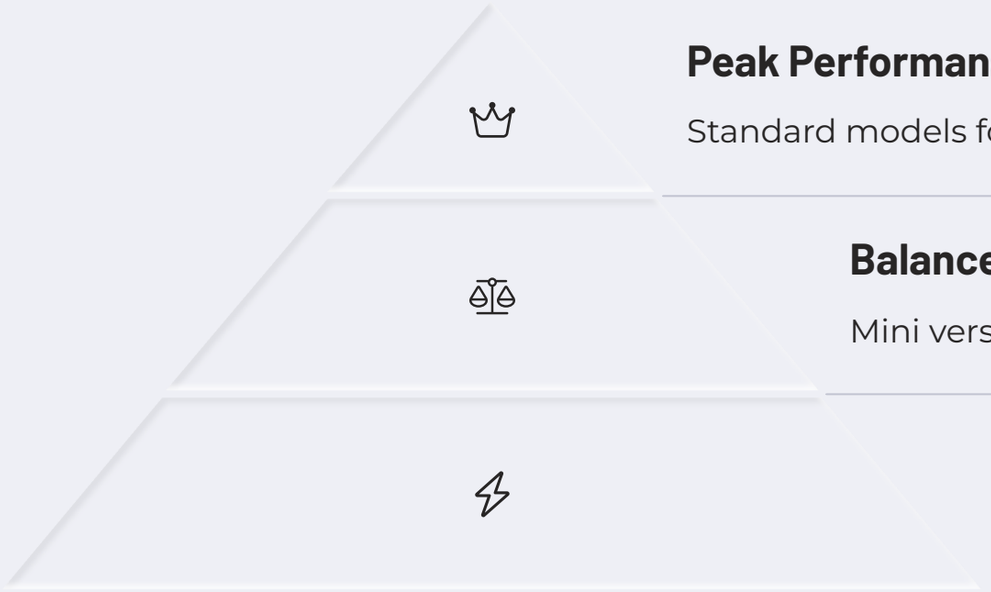
## Improved Capabilities

Better coding, instruction following

## Larger Context

Up to 1M tokens, 32k output limit

# OpenAI Strategy



### Peak Performance

Standard models for maximum capability

### Balanced Options

Mini versions for cost/performance balance

### Efficiency Focus

Nano models for speed and affordability

# Coding Breakthroughs

## 2x

### Performance Gain

Outperforms GPT-40 on Aider benchmark

## 80%

### Human Preference

GPT-4.1 web applications preferred

# AI Benchmarking

### BrowseComp

Evaluates AI agents' browsing capabilities

### PaperBench

Tests AI's ability to replicate AI research

### Standardization

Rigorous evaluation of AI reliability

# Google's Gemini Evolution

**Gemini 2.5 Flash**

Hybrid reasoning model

**Gemma 3**

Runs on single GPU/TPU

**Gemini Robotics**

Physical interaction capabilities

**Gemini Embedding**

Generalizable embeddings

# Google's Multimodal Research

### Veo 2

High-resolution text-to-video generation

### Lyria RealTime

Interactive music composition

### DolphinGemma

Decoding dolphin communication

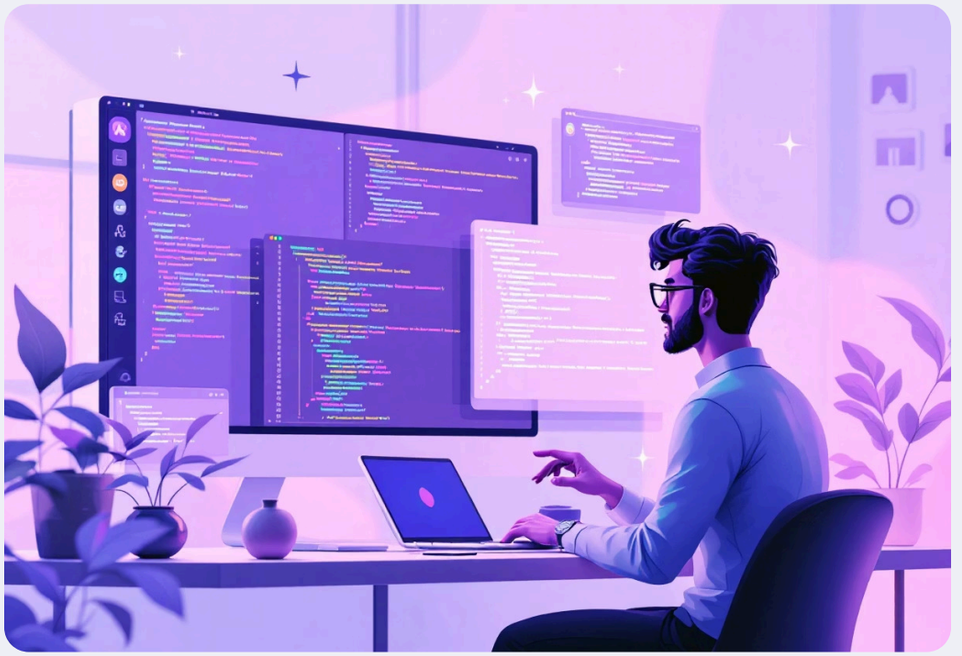
# Anthropic's Claude Advancements

## Claude 3.7 Sonnet

Most intelligent model in Anthropic's lineup


## Claude Code

Specialized model with improved coding abilities



# AI Research and Development Landscape

This presentation explores cutting-edge developments in artificial intelligence research from leading organizations including Anthropic, Meta AI, and Microsoft Research. We'll examine their approaches to model interpretability, safety, multimodal capabilities, and practical applications that are shaping the future of AI.

# Interpretability and Social Impact

## Research Focus

"Tracing the thoughts of a large language model"

"Anthropic Economic Index": AI's impact on software development

## Anthropic's Approach

Developing models while understanding their internal workings

Considering both benefits and potential disruptions to society

Demonstrating a responsible approach to AI development

# Alignment and Model Behavior

### "Claude's extended thinking"

Research exploring how Claude processes complex reasoning tasks

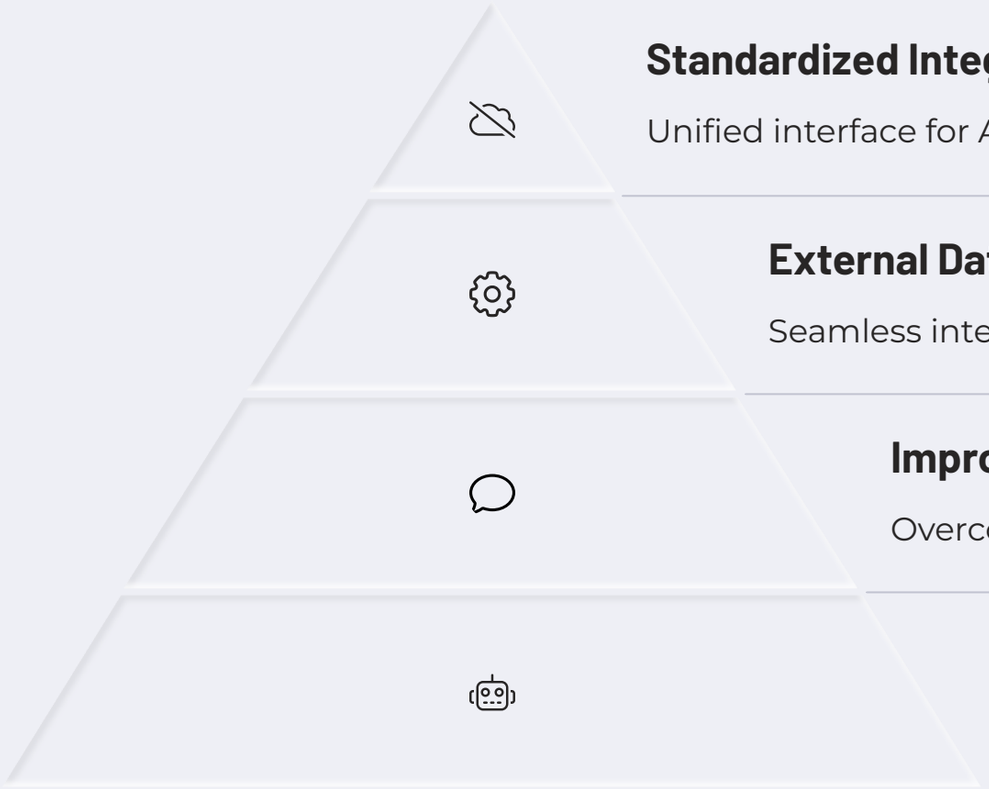
### "Alignment faking in large language models"

LLMs can feign alignment while not truly following safety guidelines

### Critical AI Safety Challenge

Ensuring AI reliability requires methods that guarantee genuine alignment

# Model Context Protocol (MCP)



### Standardized Integration

Unified interface for AI models

### External Data & Tools

Seamless interaction with various resources

### Improved Efficiency

Overcoming fragmentation of current ad hoc integrations

### Intelligent AI Agents

Facilitating creation of more capable systems

# Llama 4 Models (Scout, Maverick)

### Natively Multimodal

Designed from the ground up to process multiple types of information including text, images, and video

### Unprecedented Context Window

Up to 10M tokens, allowing analysis of vast amounts of mixed-type information

### Mixture-of-Experts Architecture

Advanced neural network design that improves efficiency and performance

### Open Source Approach

Making advanced AI capabilities available to the broader community

# Meta AI Application

### Voice-Based Conversations

Natural interaction through spoken language

### Personalized Responses

Tailored to individual user preferences and needs

### Cross-Platform Access

Available across multiple devices and services

### Widespread Accessibility

Making sophisticated AI features available to general users

# Video Generation (Movie Gen)

### Text Input

Users provide natural language descriptions of desired video content

### AI Processing

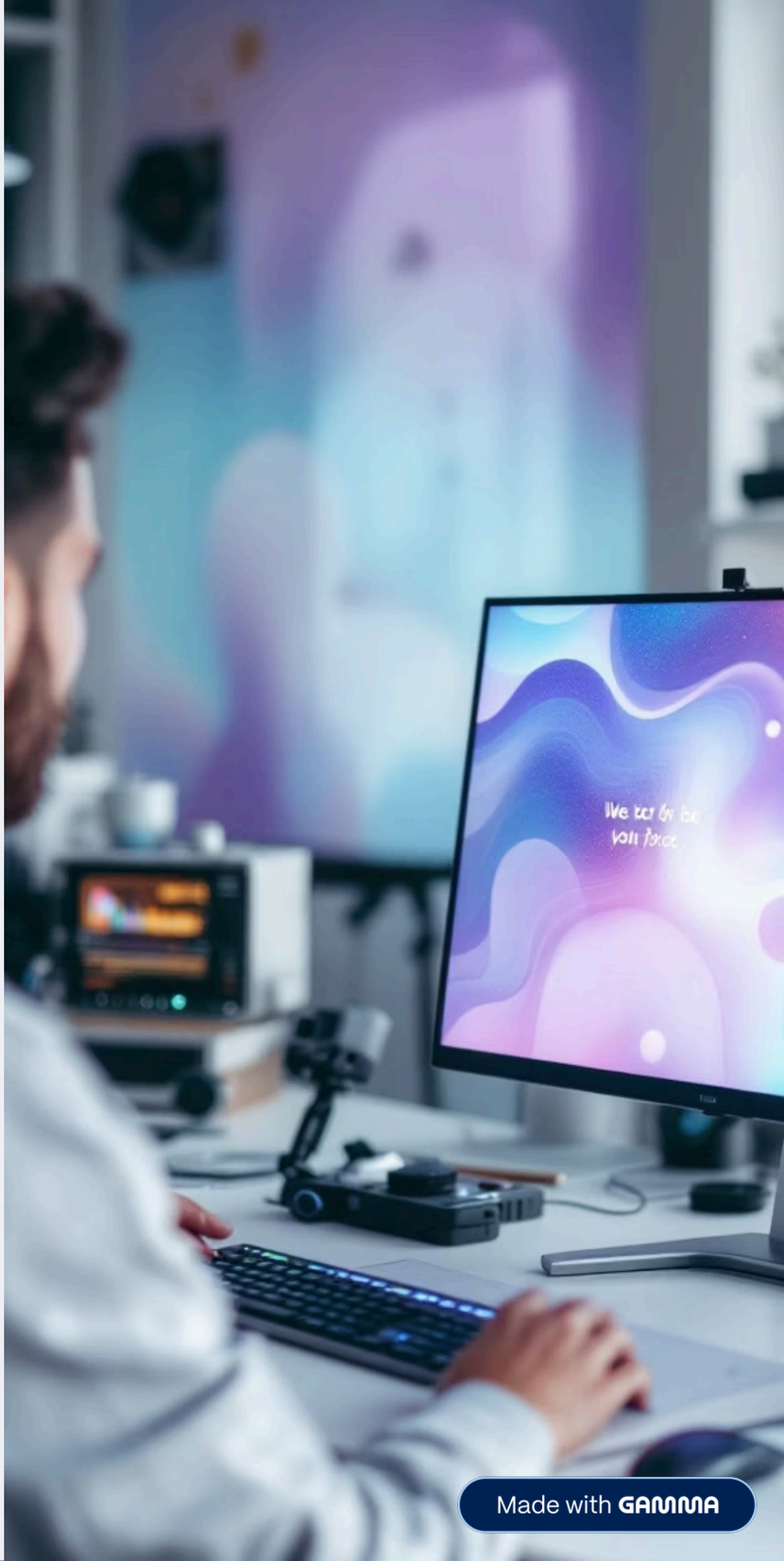
"Movie Gen" research project analyzes text and generates corresponding visual elements

### Video Creation

System produces complete videos matching the specified requirements

### Editing Capabilities

Users can manipulate and refine videos with simple instructions



# Advanced AI Research

### Llama Safety Tools

Meta AI's release of new safety tools for the open source community, building a safer AI ecosystem

### Phi-4-reasoning Model

Microsoft Research's 14B parameter model showing strong performance on complex reasoning tasks

### Efficient Model Design

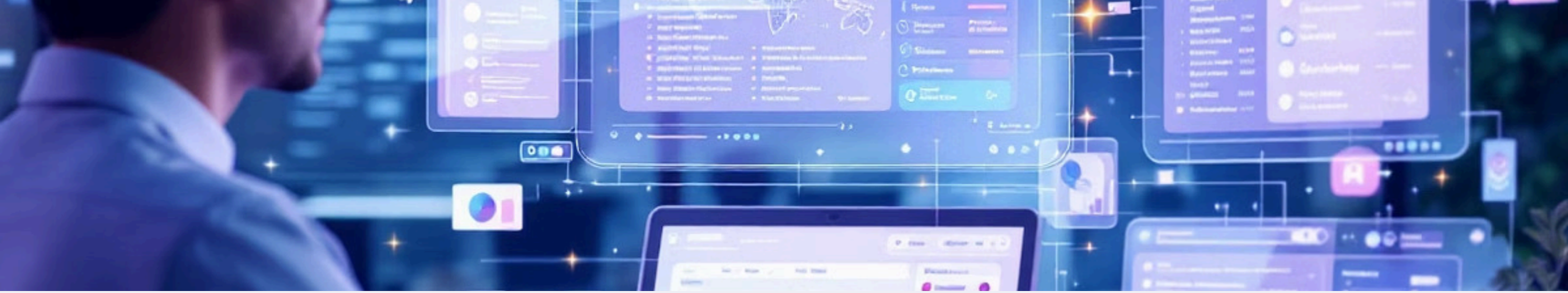
Developing reasoning capabilities in smaller models through targeted training techniques

### Responsible Development

Commitment to ethical AI advancement and collaboration with the broader community

# Recent Advances in AI Research

This presentation explores cutting-edge AI research from leading organizations including Microsoft Research, NVIDIA, Cohere, and Mistral AI. We'll examine innovations in multilingual evaluation, generative audio, novel generation paradigms, scientific applications, enterprise solutions, and open source models.



# Multilingual System Evaluation

### AI in Information Retrieval Evaluation

Study: "System comparison using automated generation of relevance judgements in multiple languages"

### Cost Reduction

LLMs can produce reliable relevance judgments in multiple languages, reducing costs

### Global Quality Assessment

LLMs show potential to automate and scale multilingual evaluation of information retrieval systems

Microsoft Research demonstrates how this can lead to more efficient methods for assessing the global quality of search engines.

# Generative Audio and Speech

### Fugatto 1

NVIDIA's introduction of generative voice converter

### Audio Large Language Models

Research: "Audio Large Language Models Can Be Descriptive Speech Quality Evaluators" (ICLR 2025)

### Unified Pre-training

Research: "UniWav: Towards Unified Pre-training for Speech Representation Learning and Generation" (ICLR 2025)

NVIDIA's research extends to generative audio. Fugatto and related research indicate growing interest in applying generative AI to audio-based tasks.

# Alternative Generative Paradigms

## Energy-Based Models

Research: "Energy-Based Diffusion Language Models for Text Generation"

These models offer advantages in dependency modeling compared to traditional approaches


## Consistency Models

Research: "Truncated Consistency Models"

NVIDIA is exploring these alternative generative models aimed at improving parallel generation and sampling efficiency

These can overcome limitations of traditional autoregressive models, offering advantages in speed

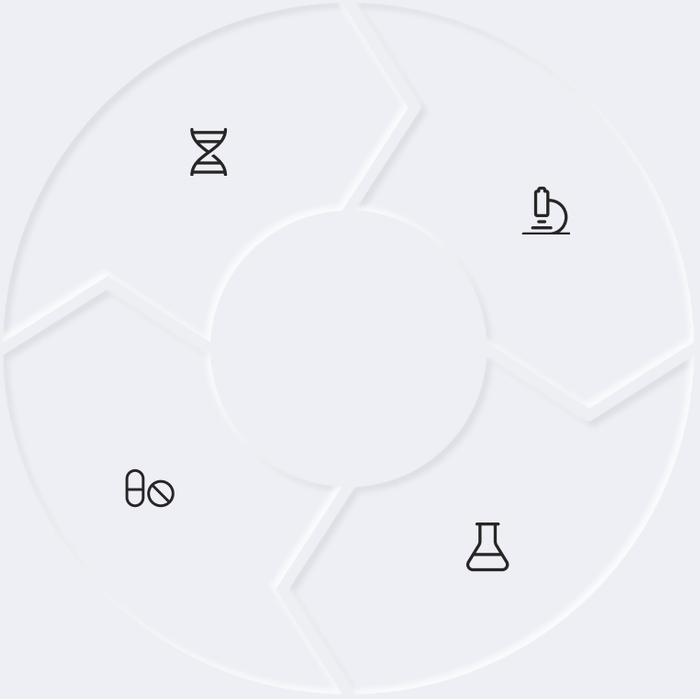
# AI in Scientific Discovery (Proteins)

### Proteina

Protein structure generation

### ProtComposer

Compositional protein structure generation



### Medical Applications

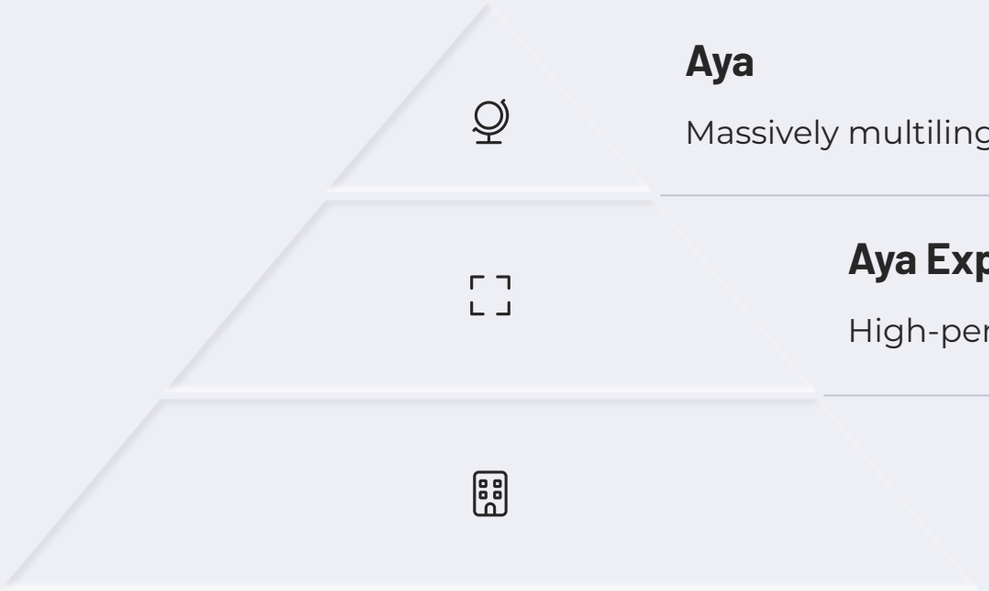
Potential for designing new therapeutic proteins

### Scientific Acceleration

Generative models accelerating biological research

NVIDIA's work demonstrates the potential of generative models to accelerate scientific research, such as designing new protein structures in biology and medicine.

# Enterprise and Multilingual Models



### Aya

Massively multilingual LLM, 101 languages

### Aya Expanse

High-performance multilingual model family, 23 languages

### Command A

Enterprise LLM for agentic tasks, multilingual

Cohere strongly focuses on multilingual AI models and tools, addressing the language gap. The Aya family shows commitment to creating AI effective across a wide linguistic range, democratizing access.

# Research Areas at Cohere

### Multilingual Research

- Multilingual LLM evaluation
- Multilingual AI safety
- Aya Vision (multimodal, accessible VLLM)

### Model Capabilities

- Vision evaluation
- Preference learning
- Coding interactions

### Responsible AI

- Model efficiency
- Fairness
- Data provenance

Cohere's research covers a broad spectrum, supporting the development of multilingual and enterprise-focused models, and addressing ethical and safety aspects of AI.

# Open Source Models from Mistral AI

**2**

### Mistral Large

Cutting-edge reasoning model

**7B**

### Mistral Parameters

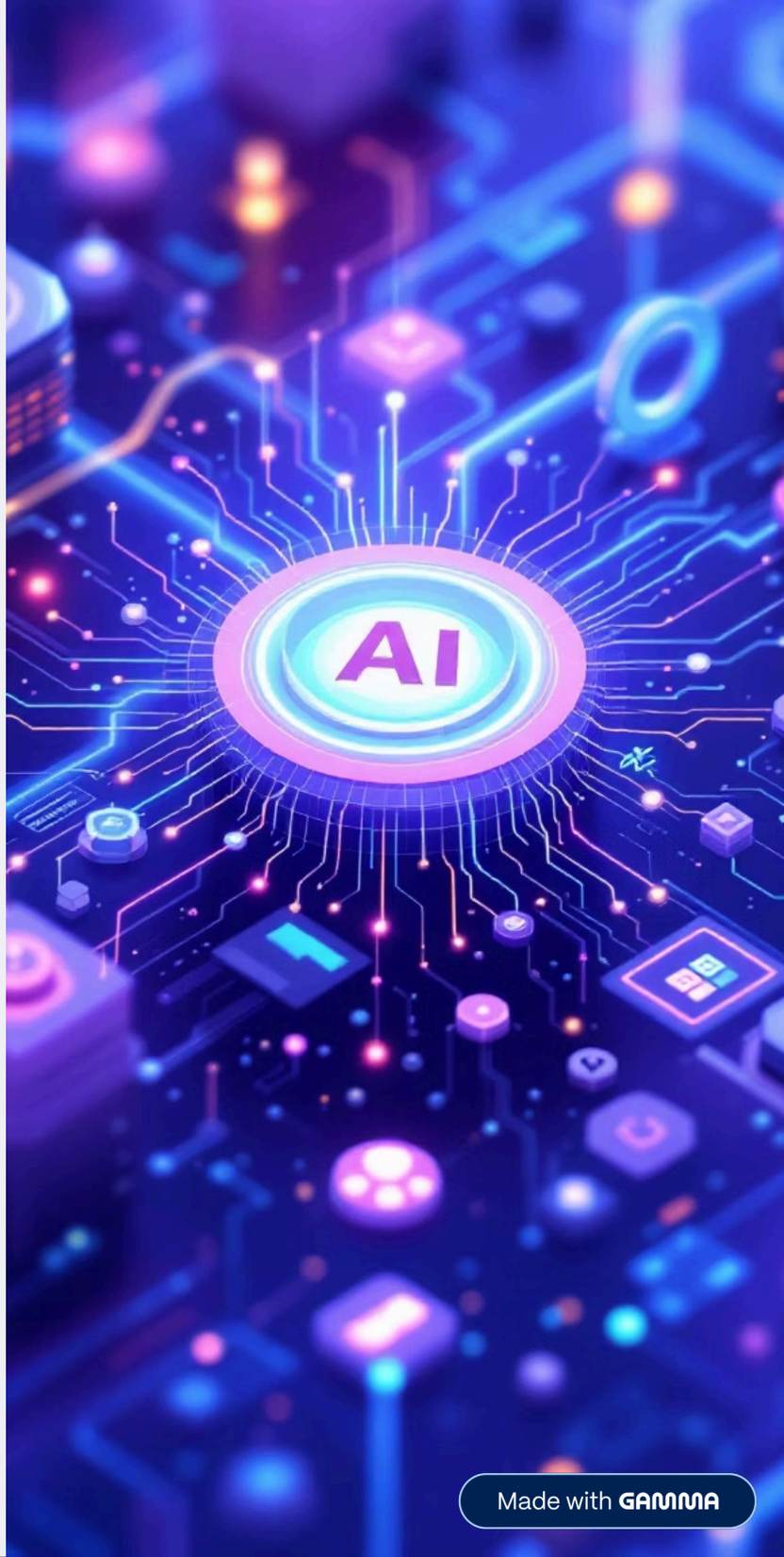
Compact yet powerful model

∞

### Innovation Potential

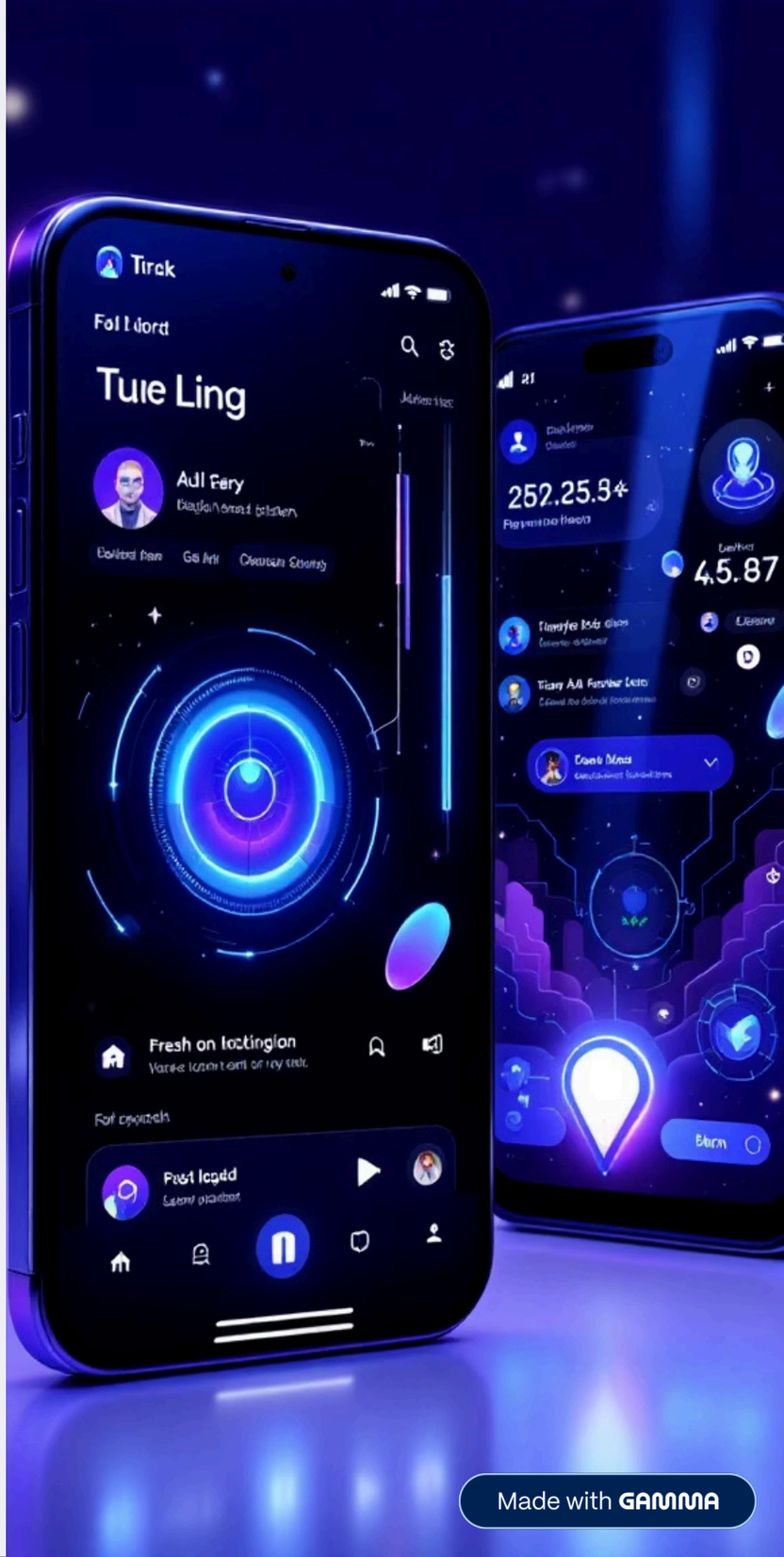
Open license fostering community development

Mistral remains a key player in the open source AI field. Releasing models under an open license allows the community to build upon and use them, fostering innovation.

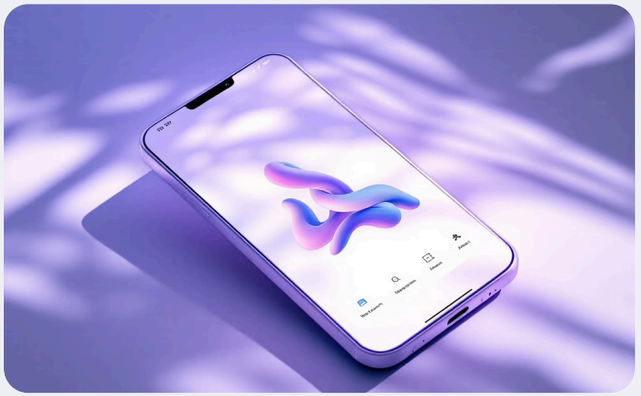


# Innovations en Intelligence Artificielle : Mistral AI, xAI et AI21 Labs

Une présentation détaillée des dernières avancées en matière d'intelligence artificielle par trois acteurs majeurs du domaine : Mistral AI, xAI et AI21 Labs. Découvrez leurs innovations récentes, leurs modèles phares et leurs contributions à l'évolution de l'IA.



# Le Chat AI Assistant



### Disponible sur iOS et Android

Lancement de l'assistant IA Le Chat sur les plateformes mobiles les plus populaires.

### Génération d'images

Capacité de créer des images directement depuis l'application mobile.


### Recherche web intégrée

Fonctionnalité de recherche web pour des réponses plus complètes et à jour.

Mistral AI vise à rendre les capacités avancées d'IA accessibles aux utilisateurs quotidiens sur les appareils mobiles. Cette initiative répond à la demande croissante d'outils d'IA disponibles en déplacement.

# Partnership with Microsoft


Un partenariat stratégique qui marque une étape importante dans l'expansion de Mistral AI sur le marché mondial de l'intelligence artificielle.

## Collaboration Stratégique

Partenariat avec Microsoft : les modèles Mistral disponibles sur le cloud Azure.

Une initiative stratégique pour élargir l'accès aux modèles Mistral grâce à l'infrastructure d'un fournisseur cloud majeur.

Augmente l'impact de Mistral et offre aux utilisateurs d'Azure un accès aux modèles d'IA européens, favorisant ainsi la concurrence.

# Grok Models (Grok 3 Beta, Grok 3 mini)

### Grok 3 Beta

Le modèle le plus avancé de xAI, avec un raisonnement exceptionnel et des connaissances étendues.

Disponible via l'API xAI pour les développeurs et les entreprises.

### Grok 3 mini

Un modèle de raisonnement rentable, offrant un bon équilibre entre performances et coût.

Solution idéale pour les applications nécessitant moins de ressources computationnelles.

La série Grok de xAI évolue rapidement, avec un fort accent sur les capacités de raisonnement. Le développement rapide et l'accent mis sur le "raisonnement à partir des premiers principes" suggèrent la création de modèles hautement intelligents.

# Grok Vision

### Analyse d'objets réels

Utilisation de l'appareil photo du smartphone pour capturer et analyser des objets du monde réel.

### Traitement multimodal

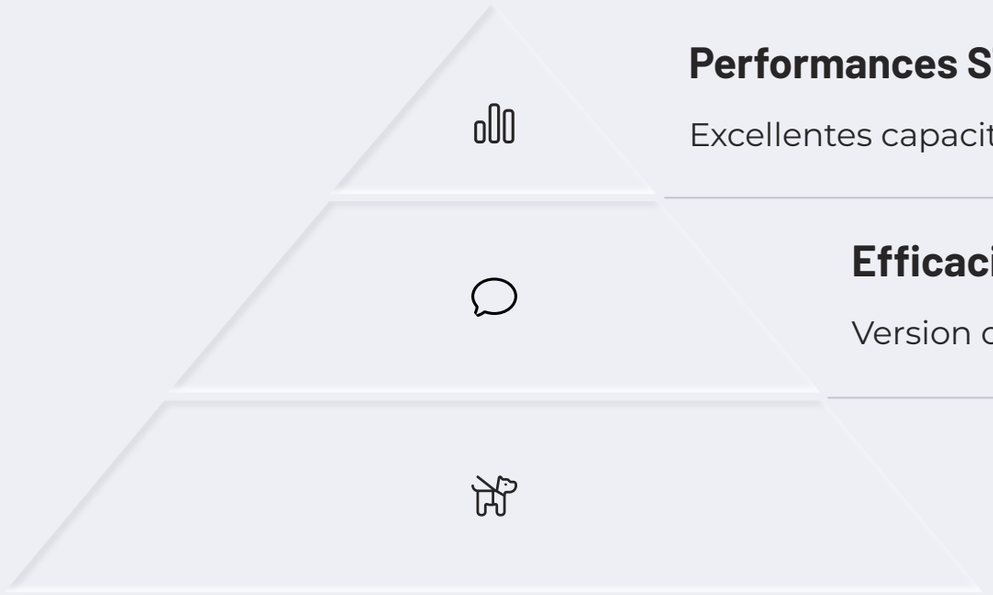
Intégration des capacités de traitement visuel aux modèles de langage existants.

### Accès premium

Fonctionnalité disponible pour les utilisateurs premium de la plateforme xAI.

Grok Vision étend les capacités des modèles Grok en traitant les informations visuelles, augmentant ainsi leur polyvalence et leur utilité dans des contextes réels.

# Cost-Effective Solutions (Grok mini)



### Performances STEM

Excellentes capacités dans les domaines scientifiques

### Efficacité

Version optimisée pour les ressources limitées

### Accessibilité

Solution économique pour un large public

L'introduction de Grok mini en tant que solution rentable vise à atteindre un plus large éventail d'utilisateurs disposant de ressources informatiques différentes. Bien que plus petit et plus efficace, ce modèle maintient de bonnes performances dans les domaines STEM.

# AI Planning and Orchestration System (AI21 Maestro)

### Planification IA

Système avancé de planification pour les tâches complexes

### Orchestration

Coordination intelligente des différentes étapes de résolution

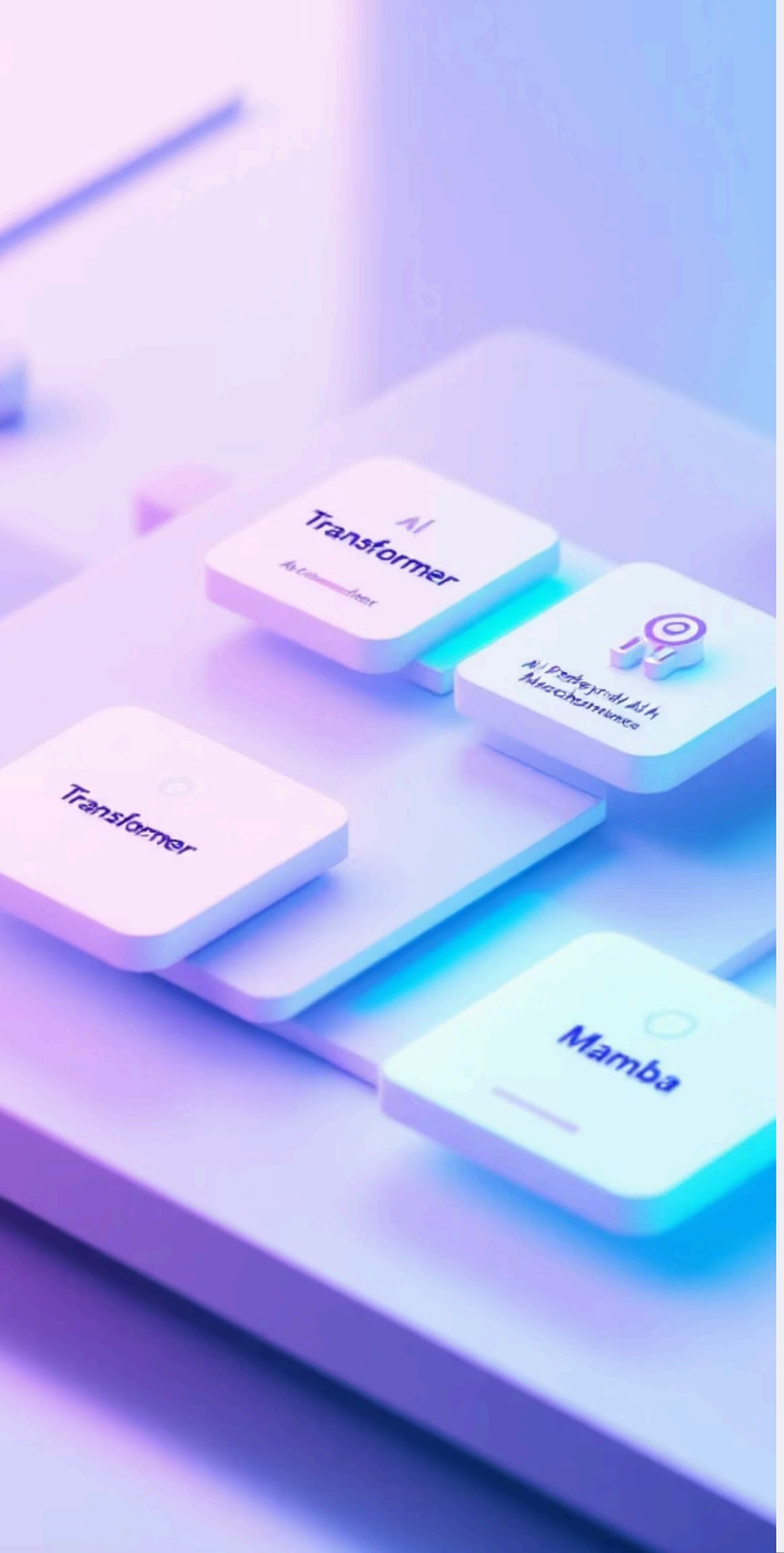
### Résolution structurée

Approche méthodique pour des solutions fiables et vérifiables

### Solutions d'entreprise

Conçu pour répondre aux besoins professionnels exigeants

AI21 Labs se concentre sur la construction de systèmes d'IA de niveau entreprise avec des capacités de planification avancées. Maestro cible des solutions d'IA commerciales fiables et vérifiables, allant au-delà de la simple génération.




# Jamba Models et Recherche en Sécurité IA

1

## Modèles Jamba

Développement d'une architecture hybride Transformer-Mamba pour un traitement efficace des contextes longs.

2

## Avantages architecturaux

Surmonte les limitations des Transformers traditionnels, notamment le coût computationnel pour les séquences longues.

3

## Recherche en sécurité

"Jamba 1.5a: Enhancing AI Safety Through Post-Post-Training Alignment"

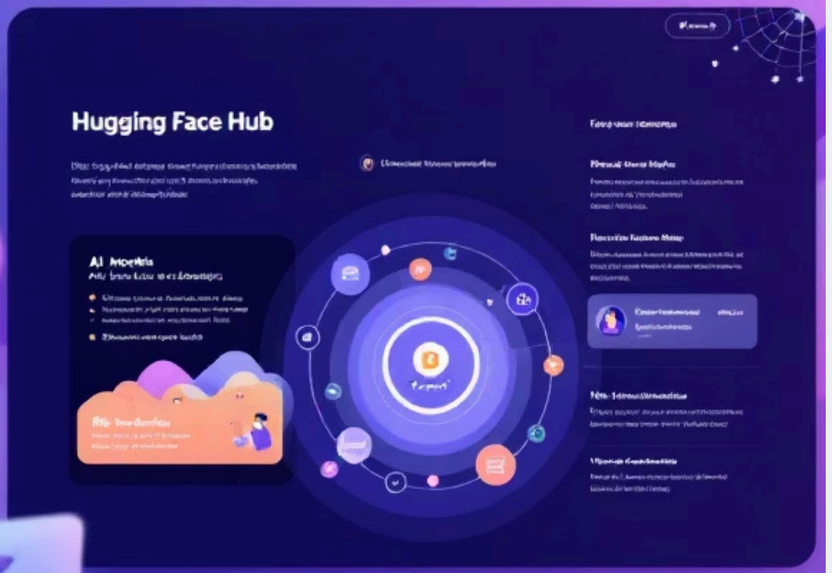
4

## Engagement pour la sécurité

Souligne l'importance de la sécurité de l'IA dans le processus de développement, même après les étapes d'alignement post-entraînement.

# Open Source AI: Community and Resources

The AI landscape is evolving rapidly with open source platforms, groundbreaking research, and ethical considerations shaping its future. Universities and organizations worldwide are contributing to advancements in AI technology while addressing its societal implications.


# Open Source Hub

### Central Platform

Central platform for open source AI models and datasets.

### Popular Models

Hosting popular models (Qwen, Dia, DeepSeek) and datasets (OpenMathReasoning, InternVL-Data).

### Key Libraries

Development of key libraries (Transformers, Diffusers, PEFT).

### Collaboration

Fostering collaboration and knowledge sharing (Hugging Face Hub).

Hugging Face and the AI community play a key role in democratizing AI technologies. The widespread adoption of the platform and libraries catalyzes innovation and lowers entry barriers.

# Reasoning Dataset Competition

### Competition Launch

Hugging Face launched a competition to create innovative reasoning datasets.

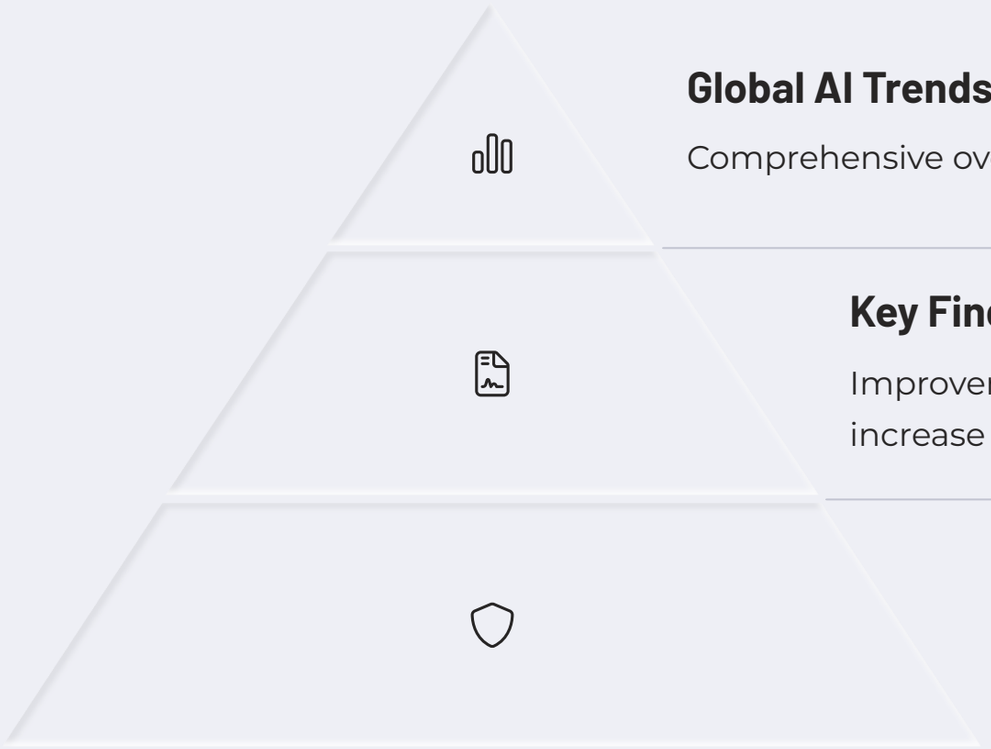
### Increasing Importance

Highlights the increasing importance of reasoning capabilities in advanced AI models.

### Initiative Goals

The initiative aims to provide the necessary data for training and evaluating these capabilities.

# AI Index Report 2025



### Global AI Trends

Comprehensive overview of global AI trends and challenges.

### Key Findings

Improvement in smaller models, decrease in inference costs, increase in AI incidents.

### Responsible AI

Emphasis on responsible AI, uneven progress in regulation and public trust.

Stanford University's Human-Centered AI Institute (HAI) provides a data-driven overview with valuable insights into the current state and future direction of AI development. The report highlights the importance of the societal aspects of the technology.

# AI Policy Research and Ethics

## Research Participation

Active participation in research and policy debates.

## AI Policy Research Symposium 2025

Presenting research on AI audit, differential privacy, AI governance, and legal use.

## UC Berkeley Tech Policy Fellows Program

Technology policy research serving the public good, focusing on AI ethics, governance, and impacts.

UC Berkeley's CITRIS Policy Lab recognizes the need for AI governance frameworks. The symposium and program topics reflect academic interest in addressing potential risks. An interdisciplinary approach is key to ensuring responsible AI.

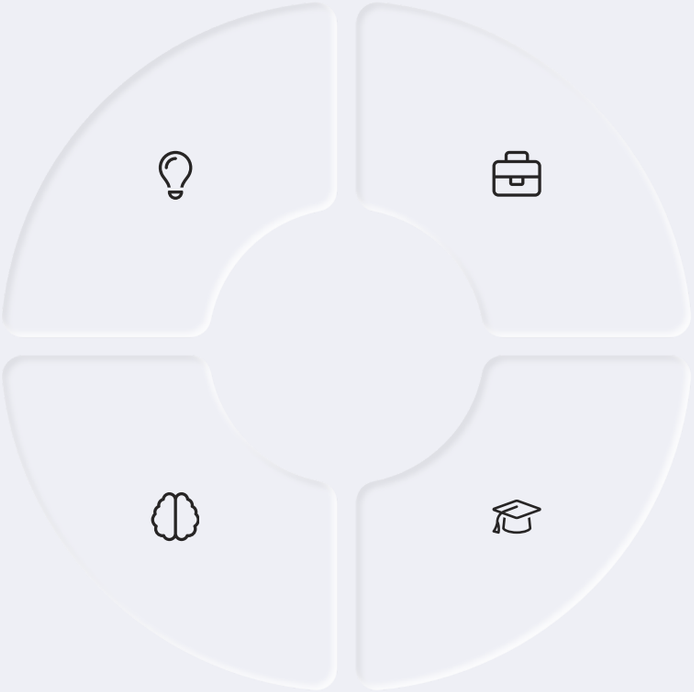
# MIT AI Conference 2025

### Knowledge

Examination of AI's impact on knowledge creation and sharing.

### Work

Analysis of how AI is transforming workplaces and job roles.



### Intelligence

Discussing the nature of artificial and human intelligence.

### Skills

Exploring the changing landscape of skills needed in an AI-driven world.

The MIT AI Conference 2025 provides a platform for discussing the broad implications of AI. Key topics include emerging trends, the future of knowledge and skills, AI in business and academia, and MIT initiatives. MIT is at the forefront of exploring the multifaceted impact of AI.

# AI and Workforce

## Complementary Role

AI is more likely to complement than replace human labor.

MIT Sloan School of Management research offers a nuanced perspective, focusing on human-machine collaboration.

## EPOCH Framework

Emphasis on human-intensive capabilities through the EPOCH framework.

This approach recognizes the unique value that humans bring to the workplace even as AI capabilities advance.

# AI Applications in Education and Healthcare

1

## AI in Education

MIT research on using AI in creating learning materials.

Highlighting benefits and the possibility of hallucinations.

2

## Deep Research Tool

MIT's tool highlights AI's potential to transform education but warns about the need to consider limitations.

3

## Healthcare Hackathon

Harvard x UofT Health Systems AI Hackathon: AI-based solutions for healthcare challenges.

Demonstrates the drive to apply AI to solve real-world healthcare problems.

# Academic-Industrial Partnerships in AI Research

Leading universities around the world are forming strategic partnerships with industry to advance AI research, while simultaneously addressing the ethical and societal implications of these technologies. This presentation explores key initiatives at the University of Toronto, University of Oxford, University of Cambridge, ETH Zurich, and EPFL that demonstrate the evolving landscape of academic-industrial collaboration in artificial intelligence.

# Industrial Collaboration (Konica Minolta)

Extension of research agreement with Konica Minolta in AI-powered sensing.

### Research Partnership

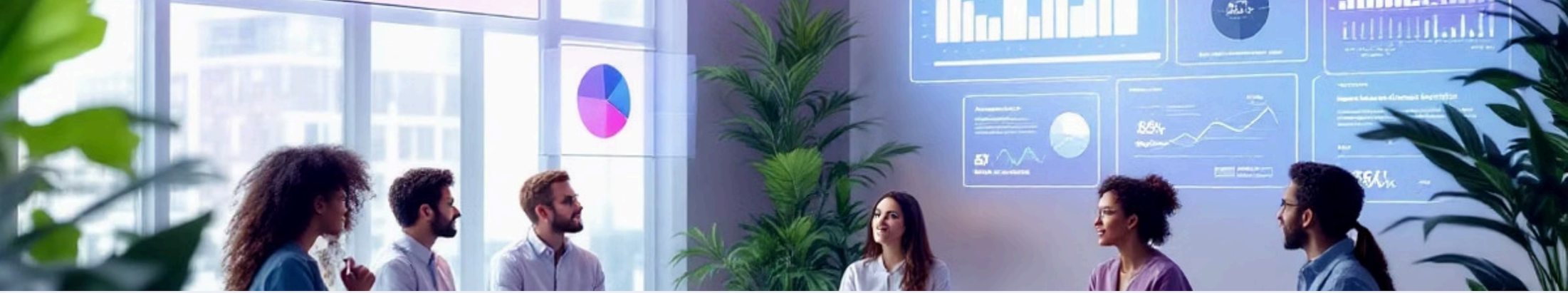
Extension of research agreement with Konica Minolta focusing on AI-powered sensing technologies

### University Strengths

Highlights the university's strength in AI-powered sensing and its ability to translate academic research into industrial applications

### Innovation Focus

Collaboration demonstrates how academic expertise can be applied to develop cutting-edge industrial solutions



# Technology and Society (Schwartz Reisman Institute)

### Research Leadership Call

Schwartz Reisman Institute call for research leaders: Impact of technology on society, AI ethics, and deployment.

### Ethical Focus

Emphasizes the university's focus on research related to the societal impact of AI.

### Interdisciplinary Approach

Bringing together experts from various fields to address complex questions about technology's role in society.

# Partnership with OpenAI

### Five-year collaboration agreement with OpenAI

Strategic partnership to enhance research and educational capabilities

### Goal: Advance AI research and education

Access to grants and AI tools to accelerate research initiatives

### AI-powered digitization of Bodleian Library material

Demonstrating AI use in knowledge preservation and accessibility

# AI Social and Ethical Research

## AI in Society Research Collection

Launch of "AI in Society" research collection examining AI's impact on economic, legal, personal, and cultural spheres

Deep commitment to exploring the ethical and societal dimensions of AI

## AI Ethics Institute Visiting Program

Dedicated program for studying ethical implications of artificial intelligence

Demonstrates a commitment to responsible development and deployment of AI technologies

# AI in Scientific Discovery

### AI for Science Summit 2024

Platform for sharing researchers' findings in AI applications for scientific discovery

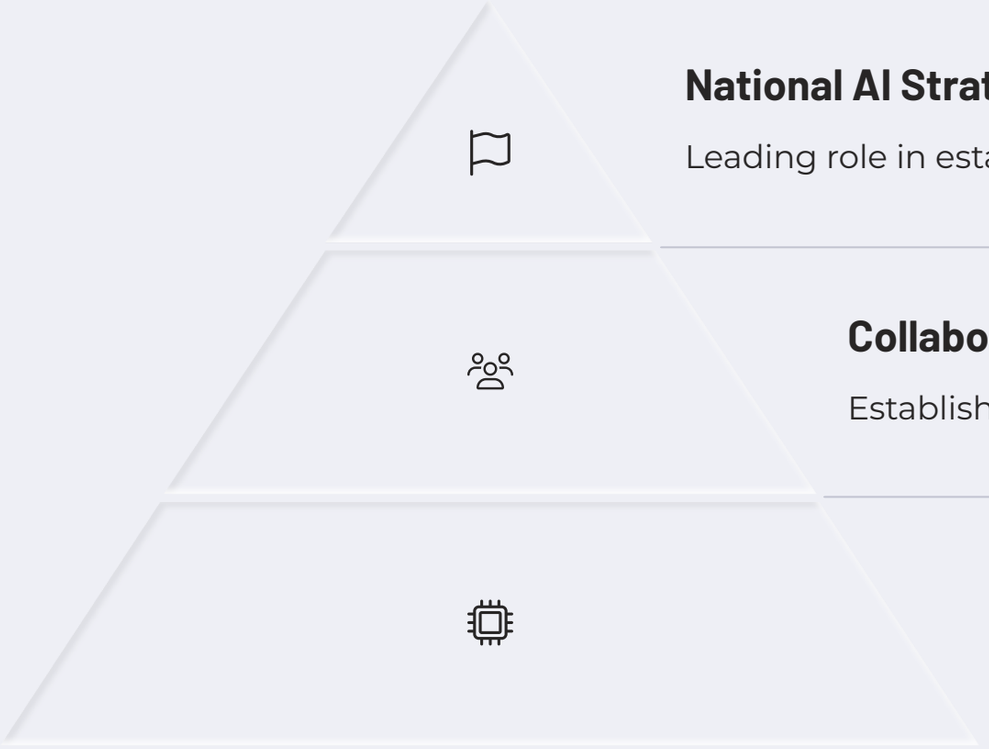
### Accelerate Programme for Scientific Discovery

Training researchers in using AI tools including Large Language Models

### AI "research clinics"

Support for scientists and students in solving AI problems in their research

# Swiss National AI Institute (SNAI)



### National AI Strategy

Leading role in establishing Switzerland as a transparent, trustworthy AI hub

### Collaborative Establishment

Establishment of SNAI with EPFL and other Swiss institutions

### Swiss AI Language Model

Develop an open, transparent Swiss AI model using Alps supercomputer

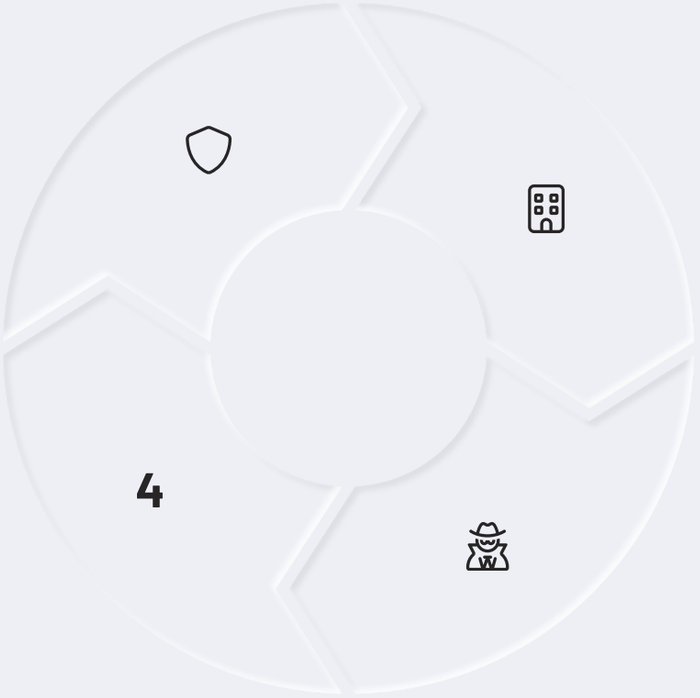
# AI Safety and Talent Development

### AI Safety Research

MARS programme connecting emerging researchers and mentors for AI safety research

### Next Generation Training

Proactive approach to meeting growing demand in software development, data science, and AI safety




### ETH AI Center

Hub for AI research and development, business-academic collaboration

### Addressing Talent Shortage

Initiatives to strengthen Switzerland as an AI hub through education

# Leading Universities in AI Research and Innovation

This presentation explores cutting-edge AI research and educational initiatives at four prestigious institutions: EPFL, Tsinghua University, Peking University, and KAIST. Each university is making significant contributions to advancing artificial intelligence through novel architectures, applications in healthcare, talent development, and foundational research.

# Brain-Inspired AI (TopoLM)

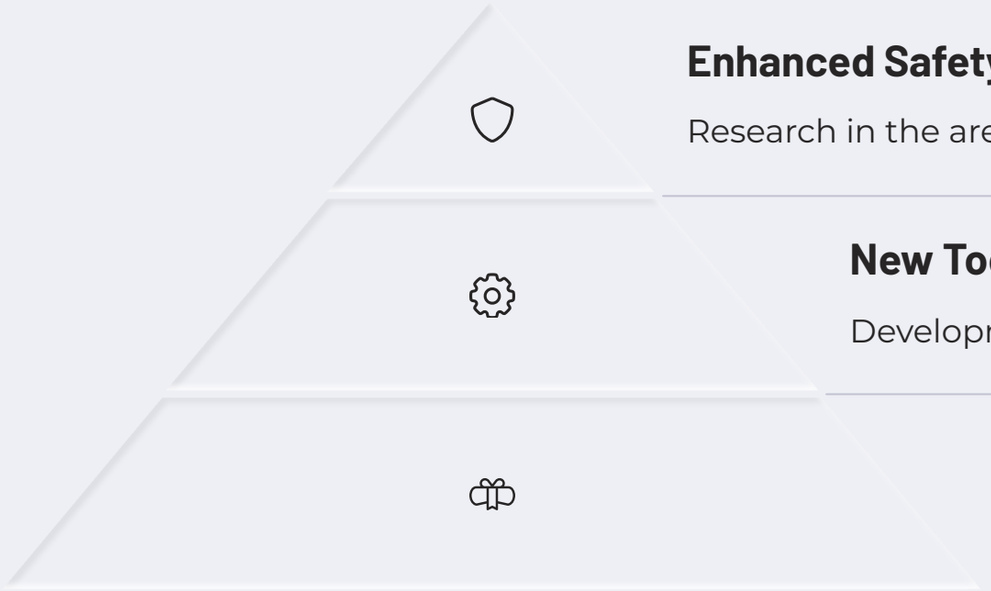
## Novel AI Architecture

Development of a brain-inspired AI language model (TopoLM).

Mimics brain organization (neuron arrangement, function).

EPFL is at the forefront of cutting-edge research, including brain-inspired models. TopoLM could represent a significant advancement in the field of brain-inspired AI.

# Trustworthy and Secure AI Tools



### Enhanced Safety

Research in the area of trustworthy AI

### New Tools

Development of new tools for building safer AI

### EPFL Leadership

Commitment to addressing potential risks

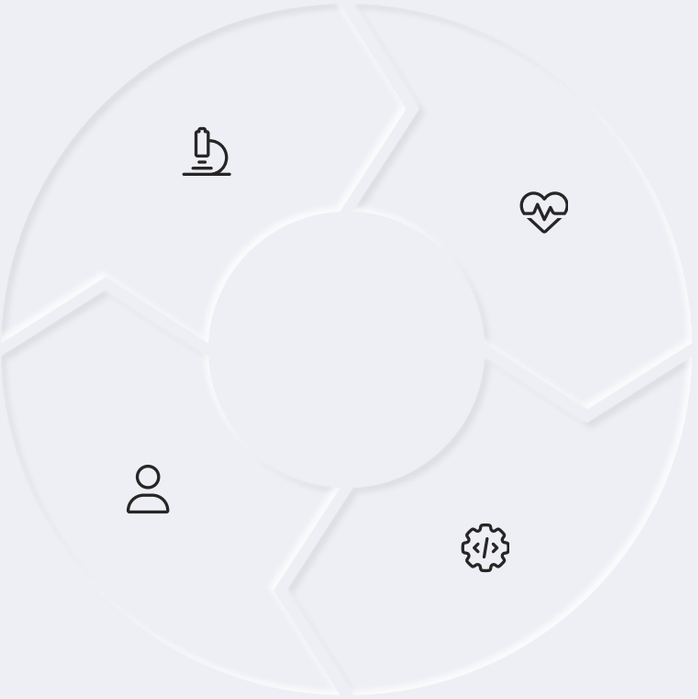
# Interdisciplinary AI Research (Fellowships)

### Science

Applying AI to scientific discovery

### Medicine

Healthcare applications



### Social Sciences

Human-centered applications

### Engineering

Technical innovations

AI Center postdoctoral fellowships for interdisciplinary AI research.

Encourages interdisciplinary collaboration. Aims to accelerate the development and application of AI techniques across various disciplines.

# AI in Medicine (AI Agent Hospital)

### AI Agents

Intelligent systems for healthcare

### Hospital Integration

Implementation in medical settings

3

### Medical Advancement

Improved patient outcomes

Launch of the Tsinghua AI Agent Hospital initiative.

Goal: Leverage AI strengths to advance medical developments.

A strategic investment in AI research and education, aiming to foster innovation in medicine. Demonstrates the university's commitment to becoming a leader in the AI field.

# AI Talent Development and Education

### New Undergraduate School

Launch of a new undergraduate school to develop AI talent.

### Interdisciplinary Integration

Emphasis on AI's role in education/research and interdisciplinary integration.

### Expanded Admission Framework

Expanding undergraduate admission framework, prioritizing the training of "AI+" talents.

Tsinghua University aims to cultivate a new generation of AI-savvy professionals. The interdisciplinary focus highlights the transformative power of AI.

# Peking University's AI Leadership

### Global Research Ranking

Leads the global institutional ranking by AI research output since 2022.

Has risen to be among the world's leading universities in AI research, showing China's growing influence.

### PKU Shenzhen Forum

2025 PKU Shenzhen Forum: Central role of innovation and AI in global economic discourse.

Actively participates in discussions about the broader implications of AI.

### Reinforcement Learning

Publication (Nature Machine Intelligence): Efficient and scalable reinforcement learning.

Presents a specific high-impact research outcome, indicating the university's deep expertise in foundational AI areas.

# KAIST's AI Initiatives

### Robotics Research

Cutting-edge development of AI-powered robotics systems

### Algorithm Development

Creating next-generation AI algorithms for real-world applications

### AI Education

Comprehensive programs developing the next generation of AI specialists

Korea Advanced Institute of Science and Technology (KAIST) continues to be a leader in AI research and education in Asia, with particular strengths in robotics, computer vision, and natural language processing.

# AI in Space Technology and Research Excellence

An exploration of specialized AI applications in space technology, research excellence in artificial intelligence, and emerging trends in the field.

# AI in Space Technology

### Development of AI-based Technique

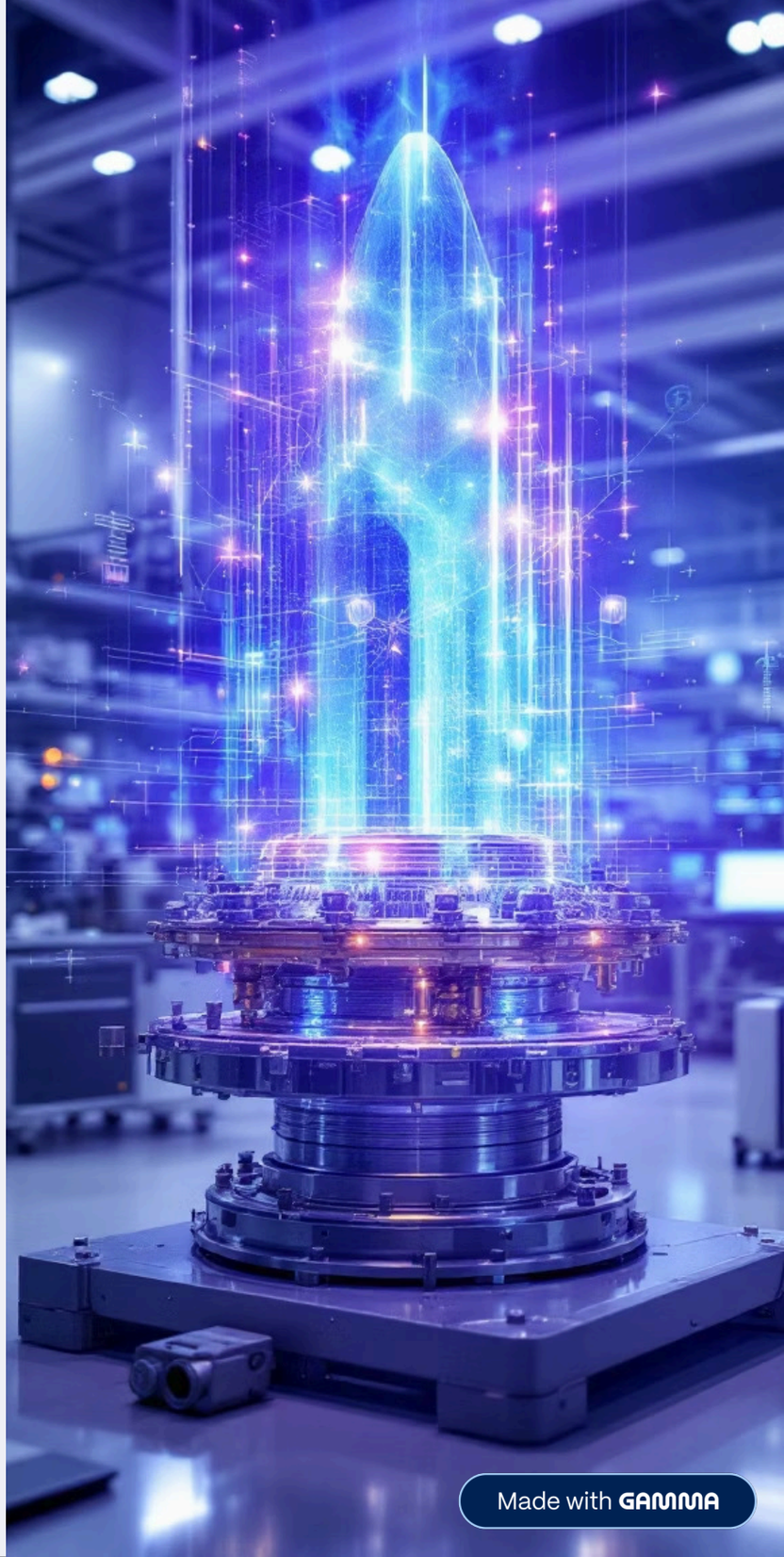
Development of an AI-based technique to predict the performance of Hall-effect thrusters (satellites, space probes).

### High-precision Prediction

High-precision prediction.

### KAIST Research

Cutting-edge research in specialized AI applications. Demonstrates AI's potential in complex engineering fields (aerospace) to accelerate design processes. Practical impact in solving real technological challenges.



# Leading AI Research Institute and Education

## Research Excellence

Ranked among the world's top AI research institutions (based on publication count).

A recognized leading institution with consistent high rankings.

## Visual AI Group Internships

Visual AI Group internship programs (generative models, foundation models, neural rendering).

Strong publication output and internship programs indicate a vibrant research community contributing to the field's advancement.

# Trend 1: Specialized AI Models

### Coding

AI models specifically designed for programming tasks

### Robotics

Specialized models for robotic control and interaction

### Biology

Models tailored for biological research and applications

### Music

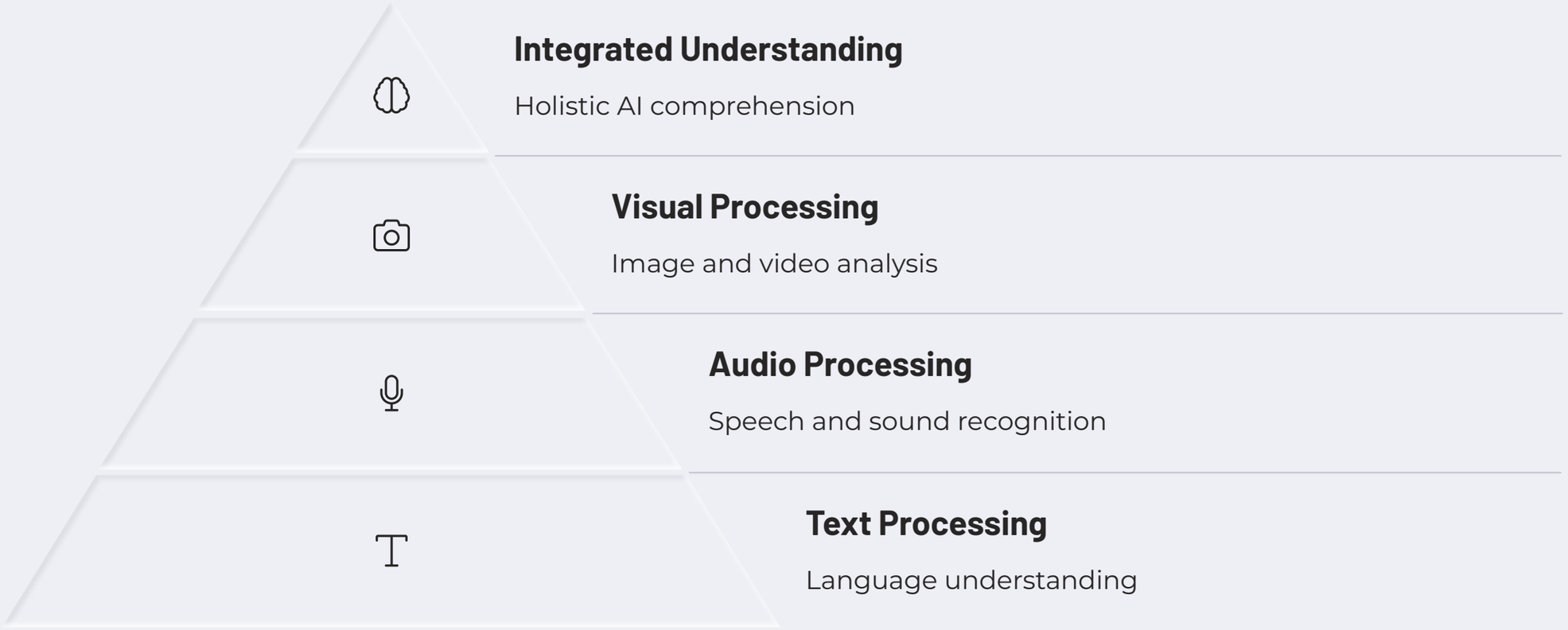
AI systems focused on music generation and analysis

Alongside general models, there is increasing emphasis on models tailored for specific tasks (coding, robotics, music, biology, geospatial, protein design).

AI is increasingly specializing to solve specific industrial or scientific problems more effectively.

Actors: Google (Gemini Robotics, Veo, Lyria), NVIDIA (Proteina), OpenAI (coding focus), etc.

# Trend 2: Emphasis on Multimodality




Growing research towards models spanning different modalities (text, image, audio, video).

Future AI systems are expected to understand the world more holistically by integrating multiple types of sensory data.

Actors: Meta AI (Llama 4), Google (Gemini, Veo), OpenAI (40 image generation), Cohere (Aya Vision), xAI (Grok Vision)

# Trend 3: Reasoning and Long Context

### Enhanced Reasoning

Improved logical problem-solving capabilities

### Expanded Context

Processing larger amounts of information

### Complex Problem Solving

Tackling more sophisticated challenges

Significant progress in LLMs' reasoning capabilities and expanding context windows.

Ability to solve more complex problems and process vast amounts of information.

Actors: OpenAI (GPT-4.1), Meta AI (Llama 4 - 10M token), Microsoft (Phi-4-reasoning), xAI (Grok 3), Google (Gemini 2.5 Flash), AI21 Labs (Jamba)

# Trend 4: Efficiency and Accessibility

### Smaller Models

Reduced computational requirements

### Consumer Hardware

Models runnable on personal devices

### Broader Adoption

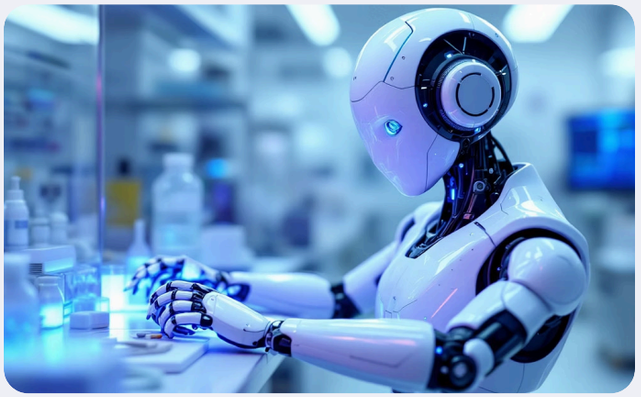
Increased accessibility worldwide

Efforts to create more efficient, smaller models (runnable on consumer hardware).

Aim is to broaden the adoption and accessibility of advanced AI technology.

Actors: OpenAI (GPT-4.1 nano/mini), Google (Gemma 3), Microsoft (Phi series), xAI (Grok mini), Mistral AI (Mistral 7B)

# The Future of AI Research



### Specialized Applications

AI systems designed for specific domains will continue to advance, solving increasingly complex problems in fields like aerospace, healthcare, and scientific research.

### Multimodal Integration

Future AI will seamlessly integrate understanding across text, images, audio, and video, creating more human-like comprehension of the world.


### Democratized Access

Smaller, more efficient models will bring advanced AI capabilities to consumer devices, expanding access beyond specialized research institutions.

# Global AI Trends and Future Directions

# Trend 5: The Growing Importance of Multilingualism

Global AI Capabilities

**Strong trend toward developing models that perform well across a broad range of languages**

**Aim to overcome AI's language barriers and ensure global access**

Players:

- **Cohere**
  Aya, Aya Expanse
- **Microsoft**
  Multilingual assessment
- **Google**
- **Mistral AI**

# Trend 6: Deeper Exploration of AI Safety and Ethics

Responsible AI Development

### Increased research to understand the inner workings of models

Ensuring safety and alignment of AI systems

### Addressing ethical and social implications

As technology advances, the need for safe and ethical applications increases

Actors:

## Industry Leaders

- Anthropic (Alignment faking, MCP)
- Meta AI (Llama protection tools)
- Google DeepMind
- AI21 Labs

## Academic Institutions

- Stanford
- UC Berkeley
- Oxford
- Cambridge
- EPFL

# Trend 7 & 8: Democratization and Integration

Spread of Open Science and Applications

### Democratization

Continuous release of open source models, datasets, libraries

### Integration

Rapid spread of applications of generative AI and LLMs in different fields

### Innovation

Open source accelerates innovation, while AI becomes a part of everyday life in more and more areas

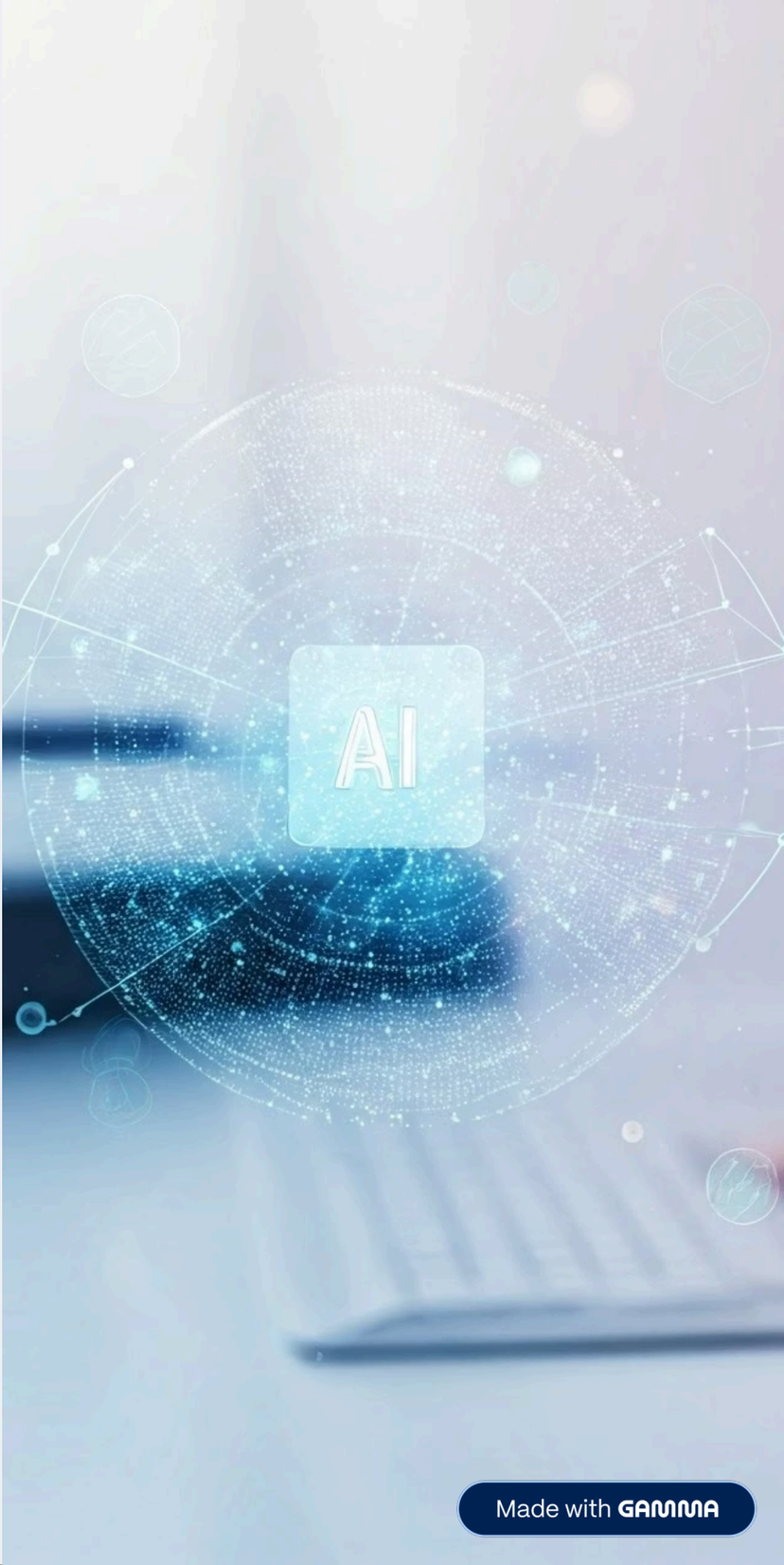
Actors:

1. **Open Source Leaders**

   Hugging Face, Meta AI (Llama), Mistral AI, Google (Gemma)

2. **Application Areas**

   Healthcare, education, science, enterprises



# Summary and Future Research Directions

### Main Results

2024-2025 brought significant advances

### Diverse Research

Performance improvement, special applications, ethical/social considerations

### Emerging Trends

Specialization, multimodality, efficiency, multilingualism indicate dynamic development

## Future Research Directions: Where to Next?

1. **Development of more robust evaluation metrics**
2. **Better understanding of properties that appear in large-scale models**
3. **Further exploration of societal impacts**

Participants: Research community

Thank you and Questions