

Open Source AI

Status, Opportunities and Challenges

Anthony J. Annunziata

Director of AI Open Innovation at IBM

AI Alliance Steering Committee Co-chair

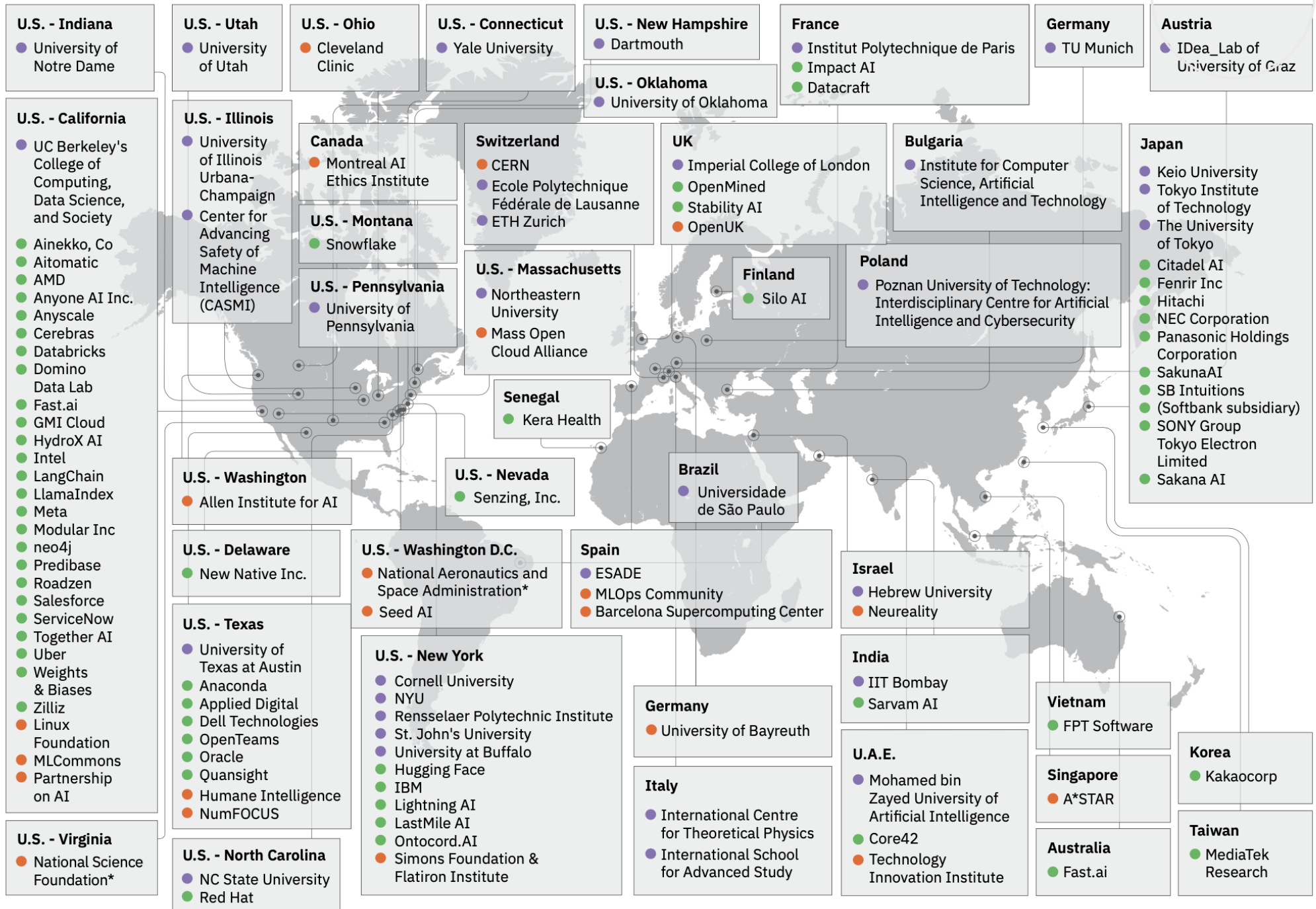
AI Alliance

A growing partnership among 117 organizations from 22 countries collaborating to advance open source AI.

Visit our booth, #129

(on the left as you enter the sponsor pavilion)

- Universities
- Startups & Enterprises
- Science Organizations & Non-profits



Five early predictions about open source AI

Five early predictions about open source AI

1. Open source AI will lag behind proprietary systems in capability. No.

Sources

<https://a16z.com/generative-ai-enterprise-2024/>
<https://www.datagravity.dev/p/open-source-vs-proprietary-models>
<https://www2.multivu.com/players/English/9240059-ibm-2023-global-ai-adoption-index-report/>
<https://ai.meta.com/blog/meta-llama-3-1/>
https://www.ey.com/en_us/services/emerging-technologies/five-ai-adoption-strategies-survey

Sources (more)

<https://lsvp.com/stories/remarkably-rapid-rollout-of-foundational-ai-models-at-the-enterprise-level-a-survey/>
<https://www.snowflake.com/data-ai-predictions/>
<https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report> Approximately 1900 IBM customer engagements in GenAI across open and proprietary model families.
Interviews and shared feedback from many of the AI Alliance's 117 members.

Five early predictions about open source AI

1. **Open source AI will lag behind proprietary systems in capability. No.**
2. **Open source AI is dangerous and closed models are safe. No.**

Sources

<https://a16z.com/generative-ai-enterprise-2024/>
<https://www.datagravity.dev/p/open-source-vs-proprietary-models>
<https://www2.multivu.com/players/English/9240059-ibm-2023-global-ai-adoption-index-report/>
<https://ai.meta.com/blog/meta-llama-3-1/>
https://www.ey.com/en_us/services/emerging-technologies/five-ai-adoption-strategies-survey

Sources (more)

<https://lsvp.com/stories/remarkably-rapid-rollout-of-foundational-ai-models-at-the-enterprise-level-a-survey/>
<https://www.snowflake.com/data-ai-predictions/>
<https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report> Approximately 1900 IBM customer engagements in GenAI across open and proprietary model families.
Interviews and shared feedback from many of the AI Alliance's 117 members.

Five early predictions about open source AI

1. **Open source AI will lag behind proprietary systems in capability. No.**
2. **Open source AI is dangerous and closed models are safe. No.**
3. **Open models will catalyze broad innovation around them (like OSS). Yes.**

Sources

<https://a16z.com/generative-ai-enterprise-2024/>
<https://www.datagravity.dev/p/open-source-vs-proprietary-models>
<https://www2.multivu.com/players/English/9240059-ibm-2023-global-ai-adoption-index-report/>
<https://ai.meta.com/blog/meta-llama-3-1/>
https://www.ey.com/en_us/services/emerging-technologies/five-ai-adoption-strategies-survey

Sources (more)

<https://lsvp.com/stories/remarkably-rapid-rollout-of-foundational-ai-models-at-the-enterprise-level-a-survey/>
<https://www.snowflake.com/data-ai-predictions/>
<https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report> Approximately 1900 IBM customer engagements in GenAI across open and proprietary model families.
Interviews and shared feedback from many of the AI Alliance's 117 members.

Five early predictions about open source AI

1. **Open source AI will lag behind proprietary systems in capability. No.**
2. **Open source AI is dangerous and closed models are safe. No.**
3. **Open models will catalyze broad innovation around them (like OSS). Yes.**
4. **Enterprises will prefer open source AI. Many do...**

Sources

<https://a16z.com/generative-ai-enterprise-2024/>
<https://www.datagravity.dev/p/open-source-vs-proprietary-models>
<https://www2.multivu.com/players/English/9240059-ibm-2023-global-ai-adoption-index-report/>
<https://ai.meta.com/blog/meta-llama-3-1/>
https://www.ey.com/en_us/services/emerging-technologies/five-ai-adoption-strategies-survey

Sources (more)

<https://lsvp.com/stories/remarkably-rapid-rollout-of-foundational-ai-models-at-the-enterprise-level-a-survey/>
<https://www.snowflake.com/data-ai-predictions/>
<https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report> Approximately 1900 IBM customer engagements in GenAI across open and proprietary model families.
Interviews and shared feedback from many of the AI Alliance's 117 members.

Five early predictions about open source AI

1. **Open source AI will lag behind proprietary systems in capability. No.**
2. **Open source AI is dangerous and closed models are safe. No.**
3. **Open models will catalyze broad innovation around them (like OSS). Yes.**
4. **Enterprises will prefer open source AI. Many do...**
5. **Enterprises will be slow to adopt OS AI until [tech, legal, policy] settles down. No.**

Sources

<https://a16z.com/generative-ai-enterprise-2024/>
<https://www.datagravity.dev/p/open-source-vs-proprietary-models>
<https://www2.multivu.com/players/English/9240059-ibm-2023-global-ai-adoption-index-report/>
<https://ai.meta.com/blog/meta-llama-3-1/>
https://www.ey.com/en_us/services/emerging-technologies/five-ai-adoption-strategies-survey

Sources (more)

<https://lsvp.com/stories/remarkably-rapid-rollout-of-foundational-ai-models-at-the-enterprise-level-a-survey/>
<https://www.snowflake.com/data-ai-predictions/>
<https://www.databricks.com/resources/ebook/mit-cio-generative-ai-report> Approximately 1900 IBM customer engagements in GenAI across open and proprietary model families.
Interviews and shared feedback from many of the AI Alliance's 117 members.

Defining Open Models



1. Steady convergence toward a practical “minimum essentials” definition of **open weight model** that includes architecture, trained weights, and documentation.

Defining Open Models



1. Steady convergence toward a practical “minimum essentials” definition of **open weight model** that includes architecture, trained weights, and documentation.
2. Broader definition of **open-source AI systems** will take time to converge and depends on developer and research community, governments and regulators.

IBM



AI Alliance

Challenges and
opportunities in
open source AI

Key Challenges – and Open Solutions

Data

Provenance, transparency, quality, safety, legality, diversity, availability for training and tuning.

Evaluation

Testing and monitoring of non-deterministic systems in various domains, especially for safety.

Domain Expertise

Infusing deep domain knowledge and skills into models and applications.

IBM Granite

A growing family of open models with data and build transparency, focused on enabling enterprise use cases.



Granite Models

Granite **Code** (3B-34B)

Granite **Language** (3B-7B)

Granite **Function Calling** (20B)

Granite **Time Series**

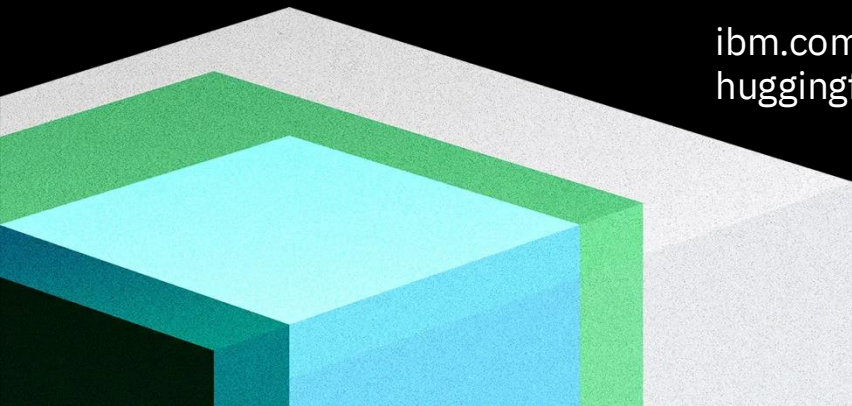
Granite **Geospatial**

Granite **Molecular**

Granite **Speculator**

Granite **Guardian**

ibm.com/granite/docs/
huggingface.co/ibm-granite



Granite Models

Granite **Code** (3B-34B)

Granite **Language** (3B-7B)

Granite **Function Calling** (20B)

Granite **Time Series**

Granite **Geospatial**

Granite **Molecular**

Granite **Speculator**

Granite **Guardian**

ibm.com/granite/docs/
huggingface.co/ibm-granite

Cooking with Granite

A community kitchen for co-creating recipes to apply Granite to enterprise use cases.



github.com/ibm-granite-community

discord.com/invite/GgDyu9jBKw



README.md



IBM Granite Community

Welcome to the Community Hub for building with IBM's [Granite Model Family](#).

The Granite family of foundation models span an increasing variety of modalities, including language, code, time series, and science (e.g., materials) - with much more to come. We're building them with transparency and with focus on fulfilling rigorous enterprise requirements that are emerging for AI. If you'd like to learn more about the models themselves and how we build them, check out [Granite Models](#).

The mission of the Granite Community organization is to work collaboratively across industries and geographies to leverage Granite to solve problems and bring value across use cases, from code generation and modernization, to forecasting and predictive maintenance, to materials discovery.

Our first goal is to build out a community kitchen for Granite, starting with cookbooks that have proven recipes to help enrich, adapt, and apply Granite models to real world applications.

The *Granite Cookbooks* provide open-source recipes for using these models in various practical scenarios.

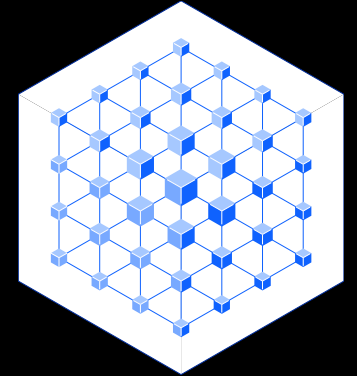
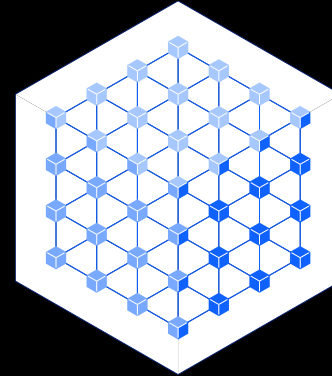
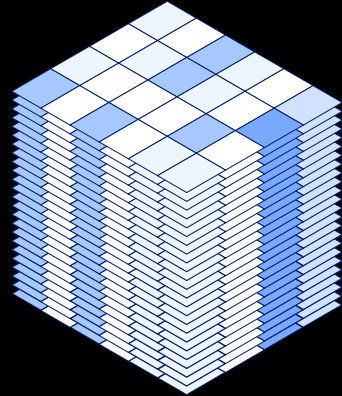
- [granite-code-cookbook](#) focuses on the [Granite Code models](#)
- [granite-timeseries-cookbook](#)

Other cookbooks are planned for materials and language.



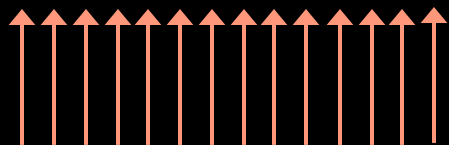
Domain Expertise: InstructLAB

Community-sourced model tuning with domain/skill taxonomies and synthetic data.



Generate examples

Knowledge sources, plus a curated taxonomy of tasks.



Community Contributions

Teacher model(s)

A teacher model generates millions of questions and answers for taxonomies.

Critic model(s)

Critic models filter the questions for correctness and quality, and prohibited content.

Student model(s)

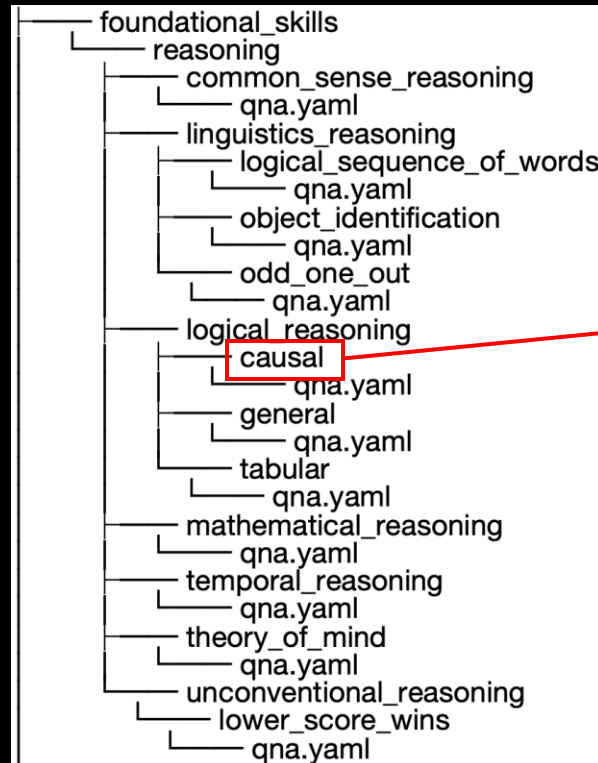
The student model is trained with the curriculum.

<https://instructlab.ai/>



InstructLAB: Models as code?

<https://github.com/instructlab>



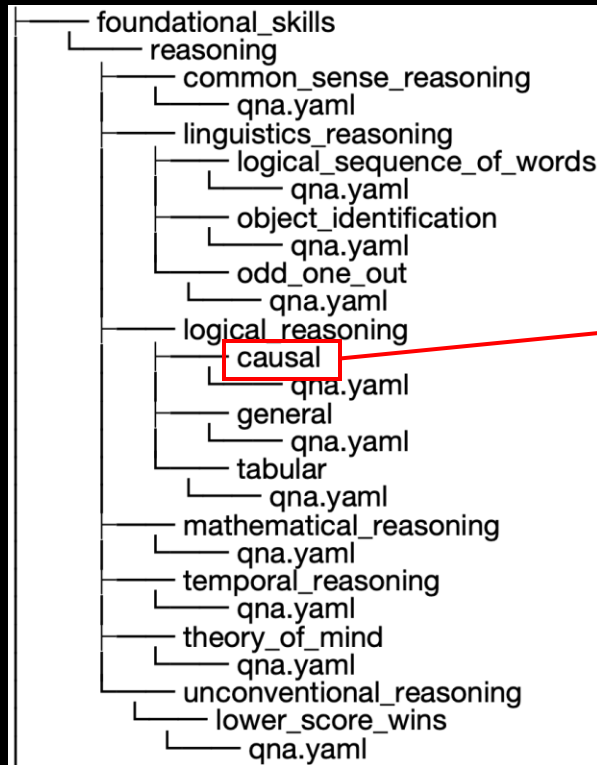
```
created_by: IBM
seed_examples:
- answer: 'While days tend to be longer in the summer, just because it is not summer doesn't mean days are necessarily shorter.'
  |
  |
  question: 'If it is summer, then the days are longer. Are the days longer if it is not summer ?'
  |
  |
- answer: 'No, we cannot conclusively conclude that some cats are black based solely on the given premises. The statement "some mammals are black" does not necessarily guarantee that among those mammals are cats.'
  |
  |
  question: 'If all cats are mammals and some mammals are black, can we conclude that some cats are black?'
- answer: 'Yes, we can conclude that all squares have four sides based on the given premises.'
  |
  |
  question: 'If all squares are rectangles and a rectangle has four sides, can we conclude that all squares have four sides?'
  |
  |
task_description: To teach a language model about Logical Reasoning - causal relationships
```

The InstructLAB project seeks to community source taxonomies of data to continually enrich model families with domain knowledge and skills.



InstructLAB: Models as code?

<https://github.com/instructlab>



```
created_by: IBM
seed_examples:
- answer: 'While days tend to be longer in the summer, just because it is not summer doesn't mean days are necessarily shorter.'
  question: 'If it is summer, then the days are longer. Are the days longer if it is not summer ?'
- answer: 'No, we cannot conclusively conclude that some cats are black based solely on the given premises. The statement "some mammals are
```

Can We Make Model Alignment More Like Software Engineering?

STRATEGY TRACK

Bay View

• **Dean Wampler**, AI Alliance and IBM Research

Models are trained by AI experts with significant compute resources, while final model alignment is often done by software engineering teams with less AI expertise, who also want model alignment to fit into conventional software engineering practices.

The InstructLAB project seeks to continually enrich model

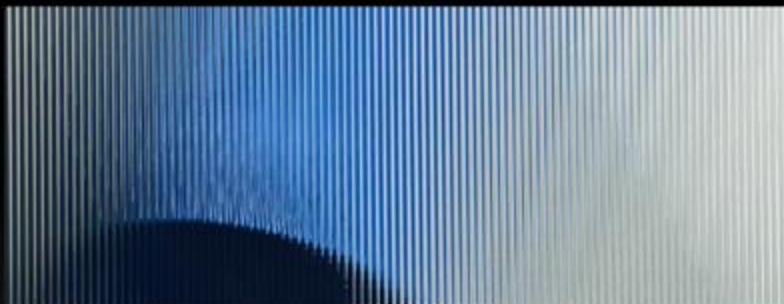
AI Alliance: Focus Areas

Advocacy



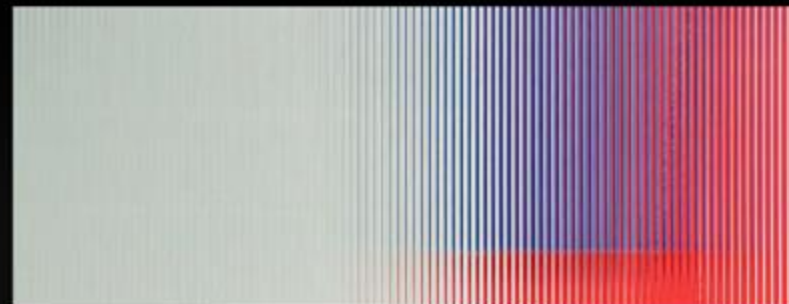
Supporting regulatory policies that create healthy, sustainable, and open ecosystems for AI.

Trust & Safety



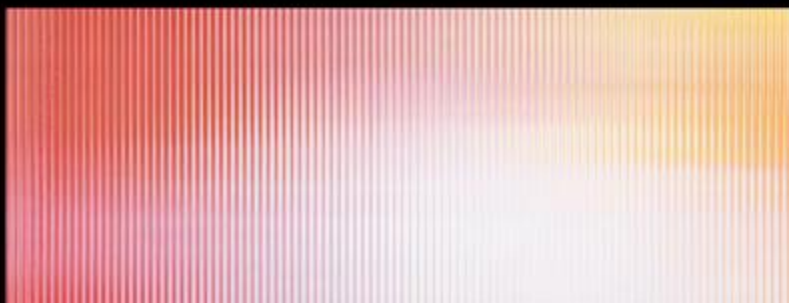
Creating benchmarks, tools, and methodologies to evaluate and ensure safe, trusted generative AI.

Applications and Tools



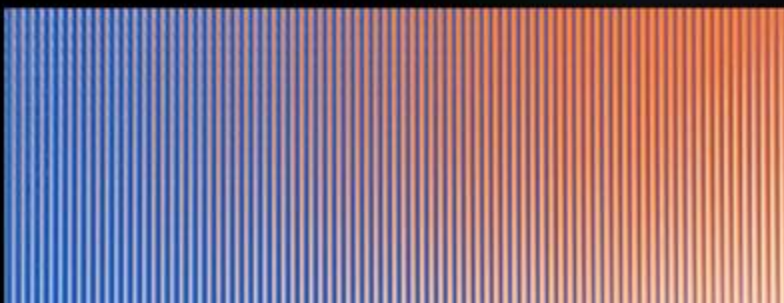
Building the most capable tools for AI model builders and GenAI application developers.

Foundation Models



Enabling an ecosystem of open foundation models, including those with multilingual and multi-modal capabilities.

Hardware Enablement



Fostering a vibrant AI hardware accelerator ecosystem through enabling software technology.

Skills & Education



Supporting global AI skill-building, education, and exploratory research.

Safety and Trust Working Group

Domain-focused Goals

Semiconductors

Climate and Sustainability

Time Series

Chemistry and Materials

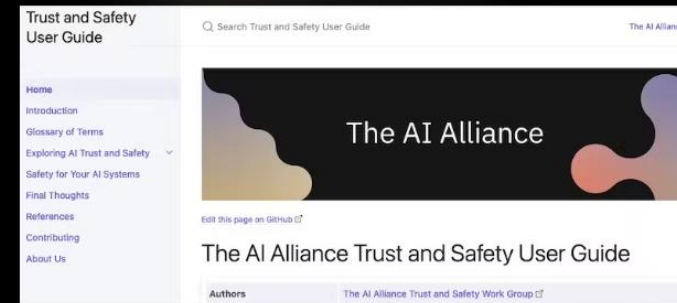
Finance

Legal

Healthcare

Understanding AI Trust and Safety: A Living Guide

Project - Trust & Safety



A major challenge for the successful use of AI is the importance of understanding potential trust and safety issues, along with their mitigation strategies. Failure to consider these issues could impact an organization's operations and the experience of its users. Concerns about safety are also a driver for current regulatory initiatives. Hence, applications built with AI must be designed and implemented with AI trust and safety in mind.

This guide offers an introduction from AI Alliance members and other experts on AI trust and safety concerns. The intent is to provide a concise introduction to the issues and offer recommendations for analyzing and mitigating these concerns.

This guide is a living document that will evolve as we broaden its coverage and incorporate new developments in trust and safety analysis and mitigation.

Click the link below to visit the online User Guide.

[Understanding AI Trust and Safety: A User Guide](#)

We welcome your contributions! The User Guide is published using GitHub Pages. Click the link below to go to the source code repository.

[GitHub source code repository for the User Guide](#)

Open Models and Data Working Group

Domain-focused Goals

Semiconductors

Climate and Sustainability

Time Series

Chemistry and Materials

Finance

Legal

Healthcare



The image shows a screenshot of the SEMİKONG website. At the top, the SEMİKONG logo is prominently displayed in white against a dark background. Below the logo, the tagline "Empowering Semiconductor Innovation" is written in white. A sub-headline reads "Meet SemiKong, the World's First Semiconductor Industry-Specific Large Language Model". A "CHAT WITH SEMİKONG" button is visible. Below this, a list of supporters includes AUTOMATIC, TEL, and AI Alliance. The main content area features the heading "Designed for Semiconductor Unique Challenges" and "TRAINED WITH SEMICONDUCTOR KNOWLEDGE". Four key points are listed with icons: 1. "Trained on a comprehensive semiconductor-related text corpus" (document icon), 2. "Novel pre-training approach leveraging domain-specific knowledge" (head with gear icon), 3. "Superior performance compared to general-purpose LLMs on industry-relevant benchmarks" (bar chart icon), and 4. "Serves as a valuable foundation for companies to build proprietary models tailored to their needs" (chip icon).

Major Initiative: Open Trusted Data

Diverse pretraining and tuning data that is open, trusted, safe and ready for AI.

Major Initiative: Industry Foundation Models

Open foundation models enriched with expert knowledge from industry verticals.

Open Trusted Data Initiative

Why?

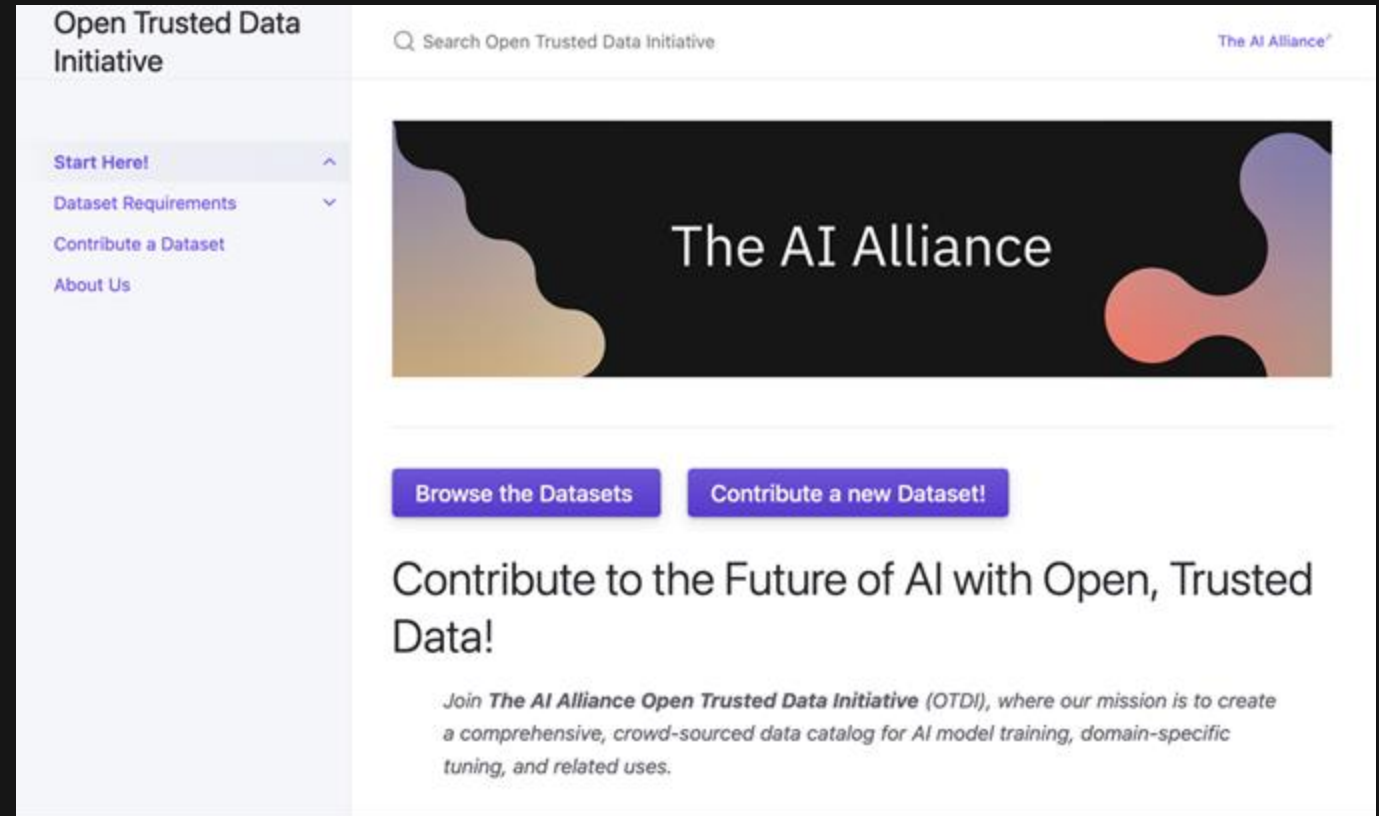
Know your data, build and deploy with confidence.

What?

Collection of carefully curated datasets for model training/tuning.

How?

Community curation and development of datasets and contributions from industry.



<https://the-ai-alliance.github.io/open-trusted-data-initiative/>

Join the AI Alliance as a collaborator



<https://thealliance.ai/become-a-collaborator>