The Ethical AI Database presents

State of the RAI Startup Ecosystem

2023 Annual Report

Table of Contents

Headlines

1. EAIDB welcomes **24 new RAI enablement startups** to the community.
2. **RAI funding remained in-line** with 2022 despite heavy VC investing headwinds that significantly reduced deal numbers in other sectors.
3. Generative AI has fundamentally changed product mixes and growth strategies for most RAI startups. **This report focuses primarily on its deep impact on the space.**
4. AI Security and Data Privacy are **currently the hottest markets** because of their direct ties with the generative AI boom.
5. The next **biggest market is Asia** - the influx of funding and technology is notable.

2H2023 Cohort

Welcoming 24 new startups to EAIDB.

Funding Landscape

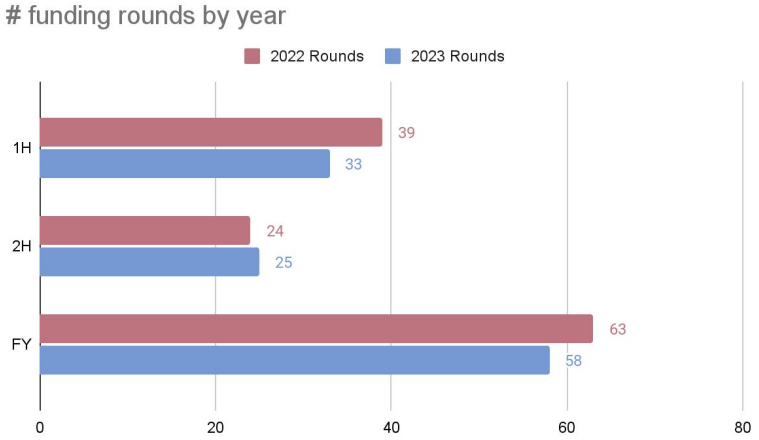
Funding: deal numbers in RAI in-line with 2022 despite stark declines in other sectors.

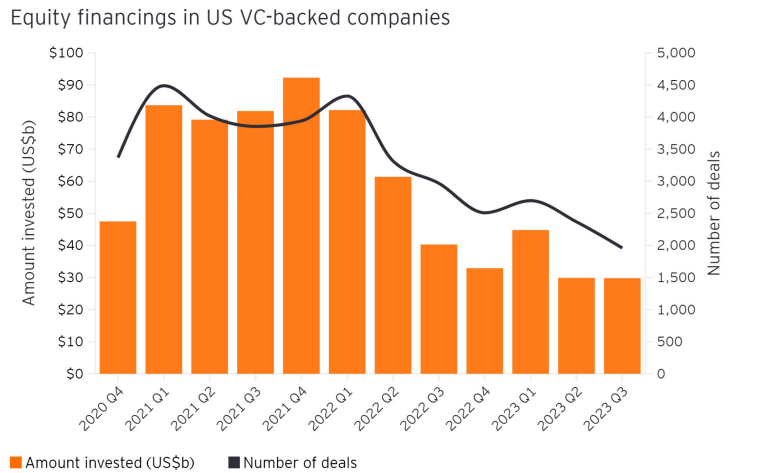
Compared to a somewhat deflated investment year due to macroeconomic and geopolitical events, RAI funding was buoyed upwards by focus on generative AI risk mitigation.

Much of the funding story for RAI in 2023 involved countering the privacy and risk downsides of generative AI. Emphasis on "AI Security" and data privacy took the lead within the RAI space as these spaces are codependent on the success of GenAI overall.

The EU AI Act has not swayed the market too much as most enterprises are still determining how to efficiently be compliant.

Deal counts for RAI in 2023 are down 7.9% vs. prior year relative to a broader investment market down 34.0% vs. prior year.





Funding information from Ernst & Young, Crunchbase

Funding: increased activity in RAI from VCs and new funds over the last few years.

HIGHEST TOTAL EXPOSURE TO RAI COMPANIES

(from inception)

TIGERGLOBAL

5 investments

PLUGANDPLAY

5 investments

Y Combinator

5 investments

FLYING FISH

4 investments

alexa fund

4 investments

NEW RAI FUNDS

moz://a ventures

3 investments

Globant

BeKind Fund

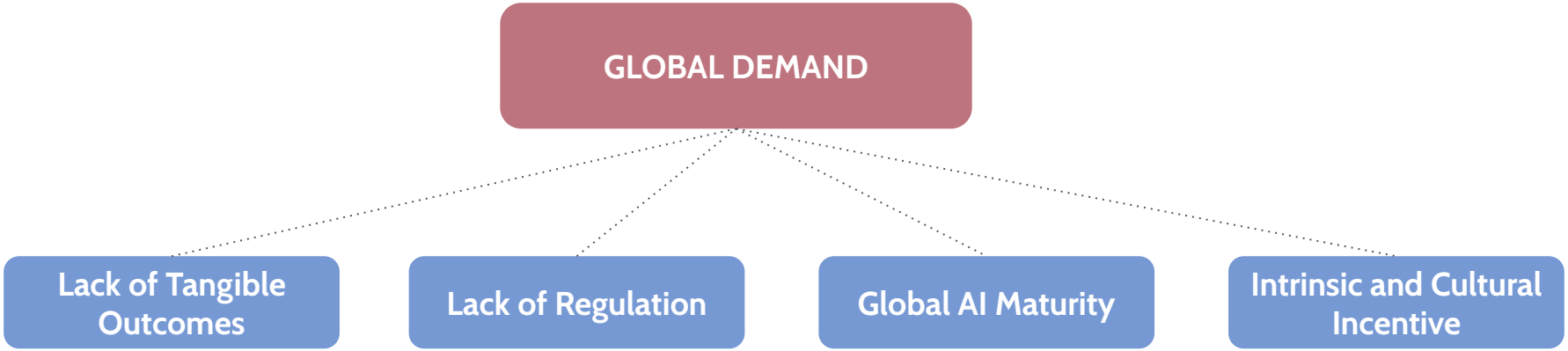
BCV

2 investments

Macro Trends

Macro: The RAI demand problem is slowly solving itself.

The four main barriers that have historically blocked significant uptake of RAI solutions have already eased from 2022 to 2023.



- AI Incident Database indicates dramatic uptick in events.
- Rite Aid, UnitedHealth, Wells Fargo, NYTimes lawsuit, etc.

- EU AI Act.
- US executive orders.
- China GenAI regulations.
- NIST AI Risk Management Framework.

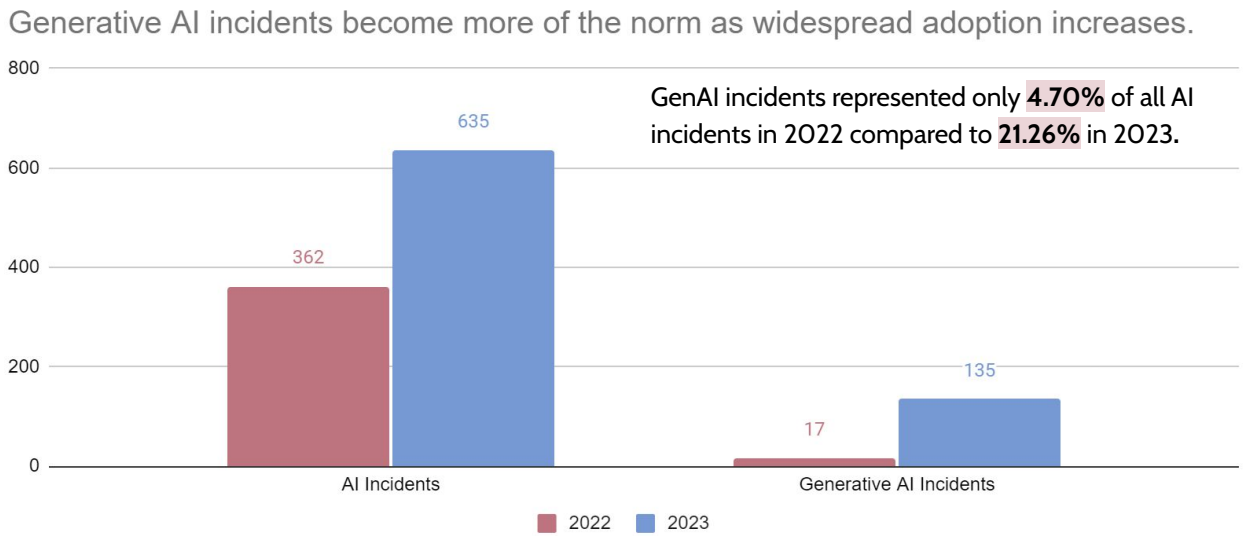
- IBM Global Adoption Index indicates steady growth.
- India, China, and UAE accelerating AI innovation.
- 82% of int'l enterprise companies surveyed have deployed or are researching AI.

- African, Nordic cultural values are rare among other nations.
- There is still very little intrinsic incentive to invest in RAI.

Macro: GenAI comes with both enormous potential and risk.

As most in the space are well-aware, generative AI comes with its own set of risks that are even more difficult to mitigate than those with machine learning and traditional AI.

According to Bloomberg, GenAI is set to become a $1.3T market by 2032. Developers and researchers are applying the same "move fast and break things" mindset that has created so much difficulty in the past. We continue to see the repercussions every day, as tracked by the AI Incident Database.



Macro: GenAI adjacent services gain significant market attention.

GenAI Security

The number of deals closed in companies providing security for AI systems (including generative AI) has **increased from four in 2022 to seven in 2023 (+75%)**. The majority of these companies were founded post-2020 and are some of the most well-established in the subsector.


MITHRIL SECURITY

CALYPSO AI


HIDDENLAYER

GenAI GRC

Of the 39 AI GRC companies in EAIDB, roughly **12 (30.8%) have already integrated one or more GenAI-specific products into their existing solutions.** The question of whether there is value in applying GRC principles to GenAI is still unsolved from an enterprise perspective - there is too much focus on sheer growth and new use cases at this point in time.


credo | ai

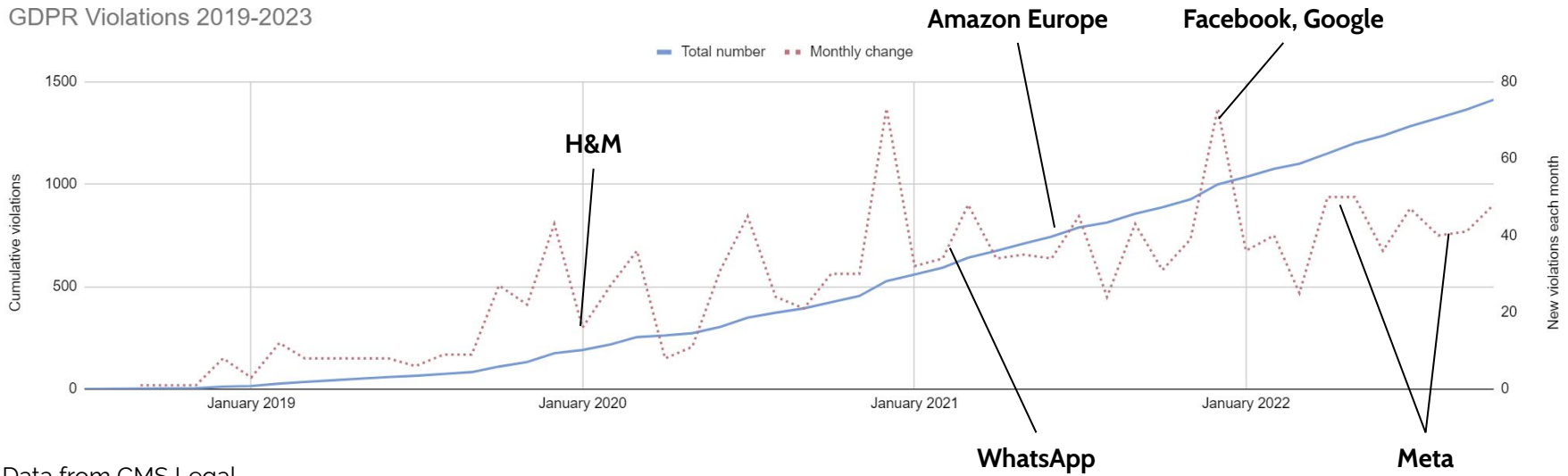
portal26


preamble

Macro: the EU AI Act will take several years before trickling down into profits.

Enforcement always lags policy.

The lag between the release of the EU AI Act and proper enforcement means **enterprises (especially those based in the United States) may not start prioritizing compliance until much later in the policy's lifecycle**. Because of the lack of enforcement, some of the most high-profile cases of GDPR violations were not found until 2-3 years after the policy's inception. It was only around this time that violations were found at an increasing rate (~40/month). Assuming a similar trajectory, **businesses may not seriously prioritize the EU AI Act until proper enforcement is enacted, which may take until 2025.**

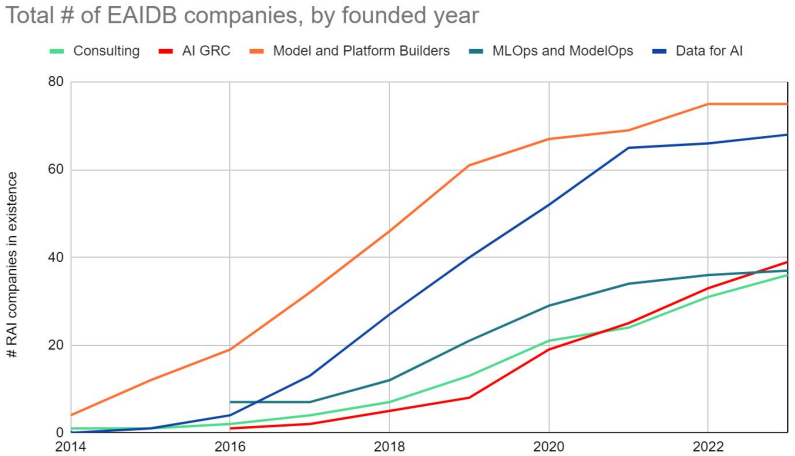


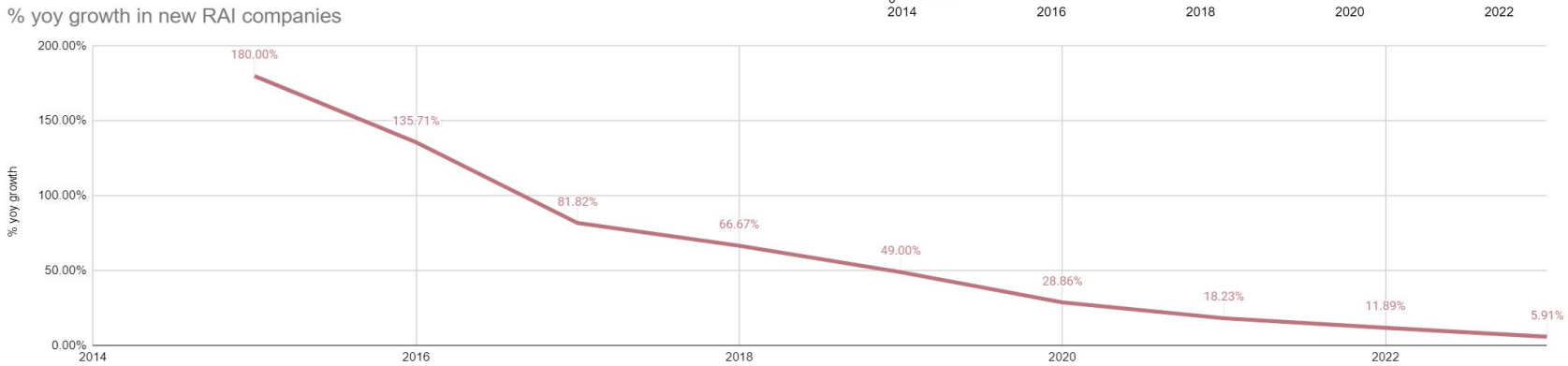
Data from CMS Legal

Macro: 2023 marked another year of slowing new entry rates.

Slow macros for RAI yield particularly challenging barriers to entry.

Combinations of factors such as general lack of real demand and the focus on AI advancement over AI trust have caused problems for new companies entering the fray. EAIDB measures another year of growth, but notes that the rate of growth has substantially reduced. **The market has reached some level of saturation** wherein products from new entrants no longer maintain significant differentiation.



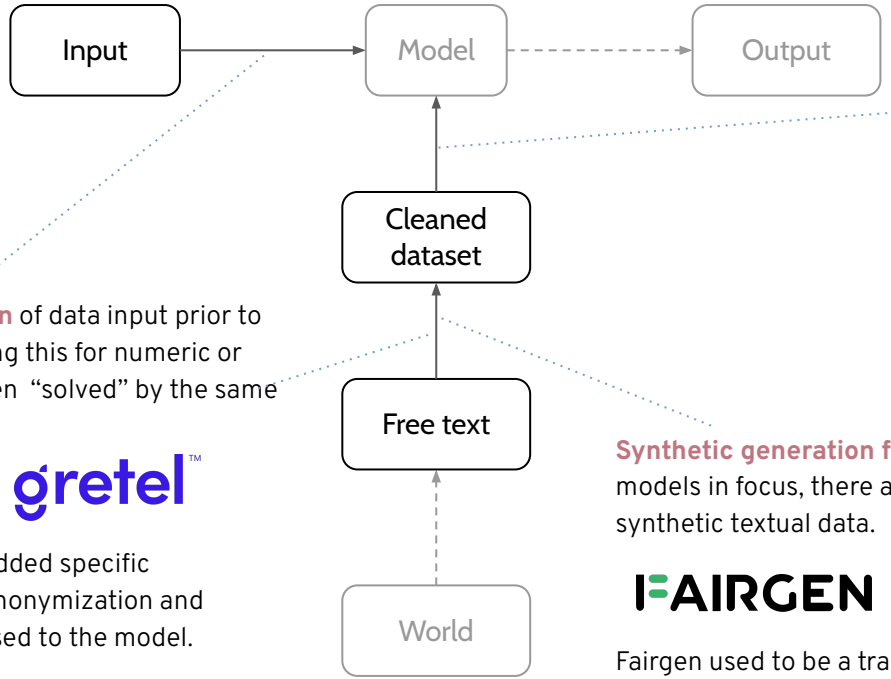


Startup Trends

Data for AI: GenAI adjacent operations boom, traditional ML concerns subside.

GenAI's inherent unpredictability and occasional inaccuracy create market opportunity.

Startups are adding more GenAI product mixes or are pivoting entirely away from traditional machine learning.



Model enhancement datasets that include reinforcement learning from human feedback (RLHF), distillation, or other methods can dramatically improve performance of chat LLMs.

Snorkel

Humans in the Loop

These companies started off as traditional sourcing and labeling companies but have added GenAI-related product mixes.

Anonymization or **redaction** of data input prior to inference or finetuning. Doing this for numeric or tabular data has already been "solved" by the same companies.

PRIVATEAI

gretel

Private AI and Gretel have added specific capabilities to handle text anonymization and intercept PII before it is passed to the model.

Synthetic generation for finetuning. Given that LLMs are now the models in focus, there are new opportunities here specifically for synthetic textual data.

FAIRGEN

TONIC

Fairgen used to be a traditional ML debiasing company before pivoting into GenAI.

Data for AI: increased focus on data management as finetuning becomes the norm.

How do enterprises ensure that sensitive data is not used to train models?

1. Enterprise-wide data governance, mapping, and PII detection (*passive management*)

1touch.io BigID Transcend harmonic

2. Redaction, obfuscation, and synthetic replacement (*active management*)

PRIVATEAI hazy

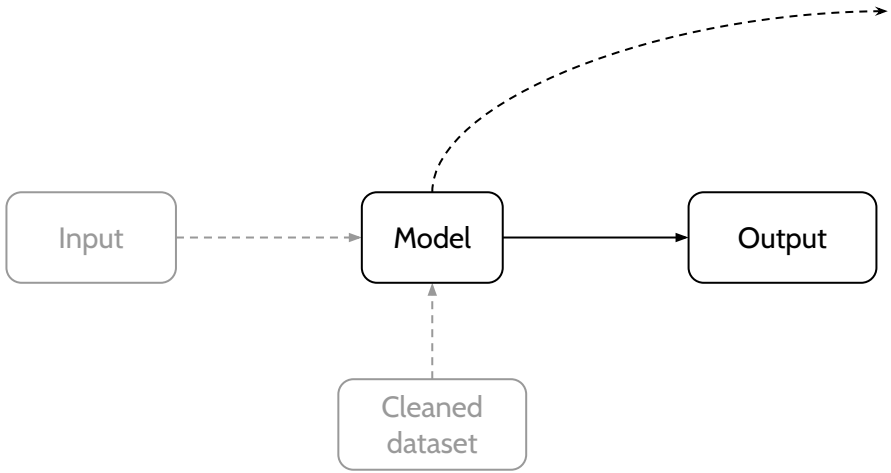
3. Privacy-preserving AI and other model-based methods (*model and pre-model management*)

inpher DynamoFL TripleBlind™

Many passive management systems also leverage active management methods to resolve any conflicts, issues, or leakages.

MLOps and ModelOps: added product mixes but no major pivots to GenAI.

Besides adding a few tools to their product mix, MLOps tools remain roughly unchanged from 2022.



Pre-deployment

- Testing and validation.
- Debiasing.
- Explainability.

Post-deployment

- Bias and drift monitoring.
- (GenAI) feedback looping and finetuning.

kolena

modzy™

ZETANE

trubrics™

These companies are benefiting from widespread and accelerated adoption of AI, which will continue as a trend in 2024.

MLOps and ModelOps: Asia is scaling their AI operations with unparalleled speed.

Asia sees new MLOps companies, successful US/EU MLOps companies expand abroad with both traditional AI and GenAI in mind.

- KKR, Bain Capital invest in data center operations in Asia (Bloomberg, paywall).
- Fiddler AI raises funding from Dentsu Ventures to expand in Japan (Dentsu).
- HuggingFace's leaderboard is routinely dominated by Asian developers and research institutions (HuggingFace).
    - SusChat-34B from Southern University of Science and Technology and IDEA-CCNL.
    - TigerBot from Tiger Research.
    - Yi series from 01.AI.

Asian RAI MLOps companies

DeepBrainz

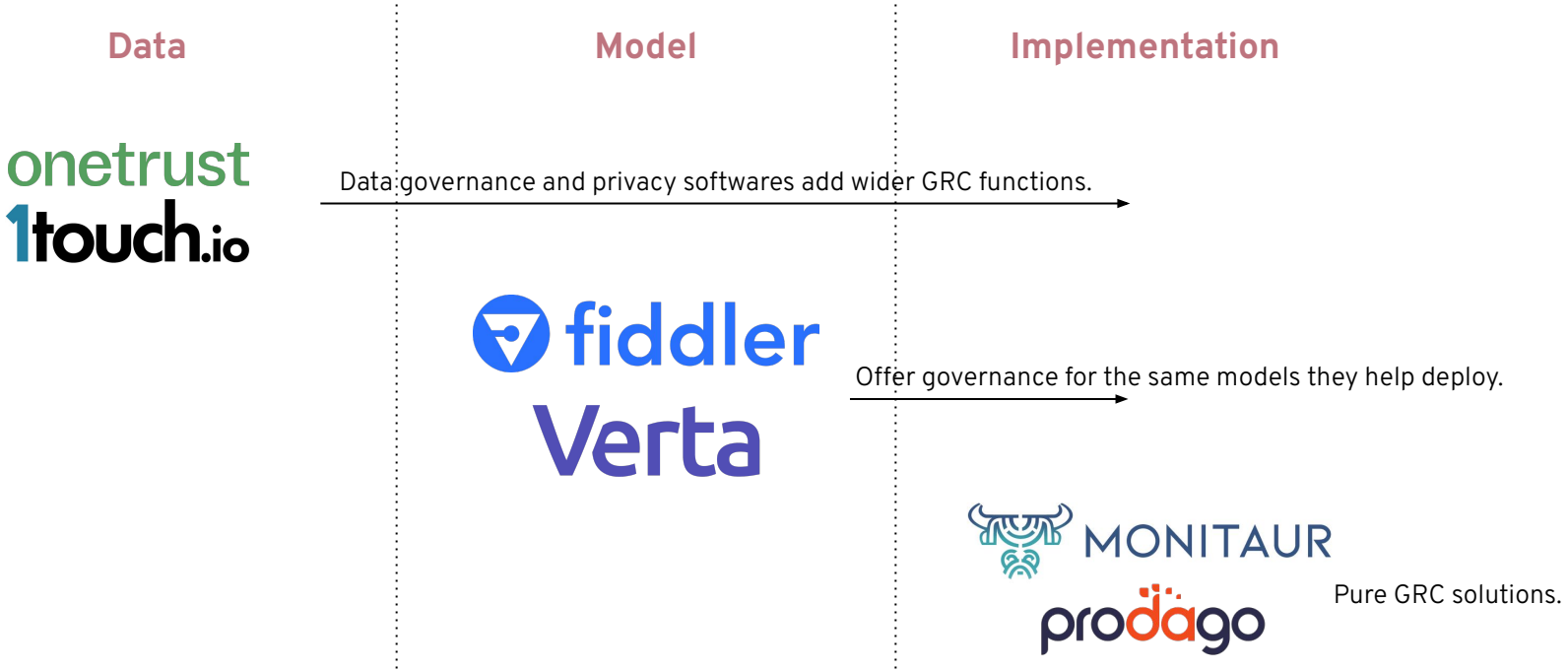
XAI by Arya.ai

akira AI

AI GRC: a growing market faces increasing competition.

RAI startups recognize enterprise need to comply with new legislation.

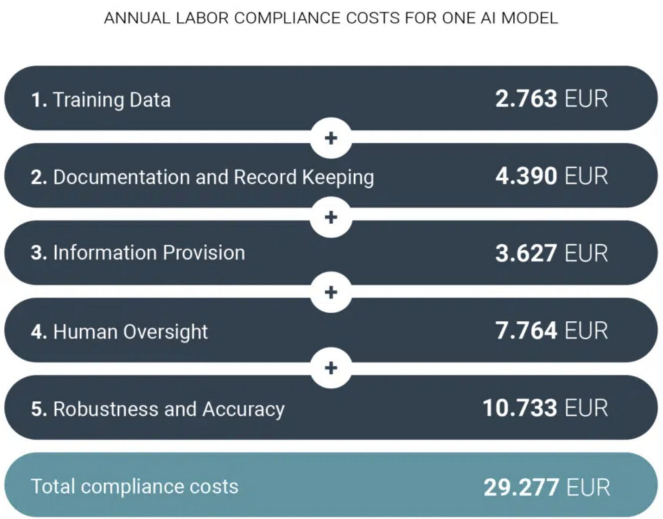
Only a few years ago, GRC was an underperforming sector because of i) lack of demand and ii) lack of tangible outcomes. Now, they face improving demand but worsening competition as startups from other sectors adopt GRC as a pillar.



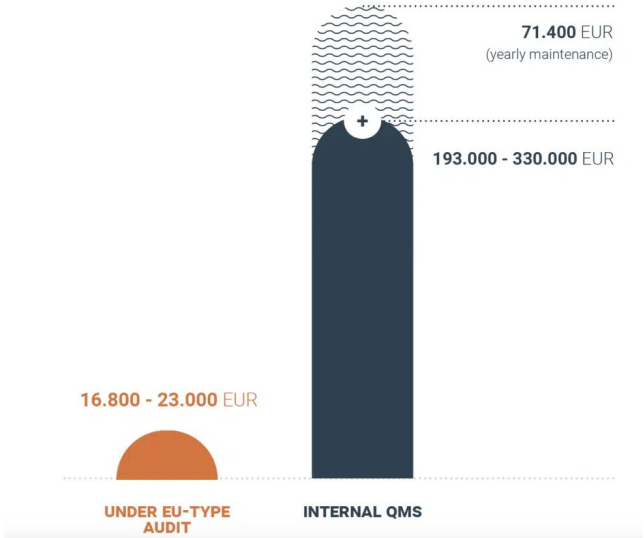
AI GRC: the cost of compliance is a key driver for demand.

Compliance is expensive – almost 30,000 EUR per model.

An article written by 2021.AI (an EAIDB company!) discusses the various costs related to compliance with the EU AI Act. These consist of a **labor cost (variable)** and a **certification cost (fixed)**, since the system itself must be certified either by an EU-type audit or an internal quality management system. **AI GRC startups are promising cost savings on several of these steps.**



AI GRC companies promise cost-savings on these four steps and allow external audits to be much more efficient.

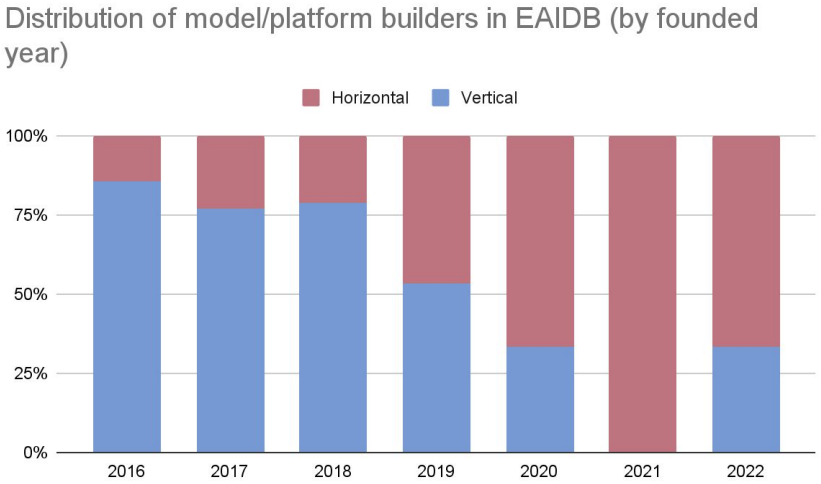


Model/Platform Builders: horizontal scaling is the favored approach.

Limited early use cases for GenAI outside of select verticals like healthcare, military, etc. creates challenging environments for vertical builders.

The past few years have seen declining numbers of new, vertical-specific companies and increasing numbers of wider-market business models.

This may have been in part due to the extremely limited demand for responsible AI enablement in the past. It may be that, going forward, **more verticals will require vertical solutions.**



Responsibly designed LLMs in medical settings

MABEL

mpathic

Responsibly designed LLMs in general settings

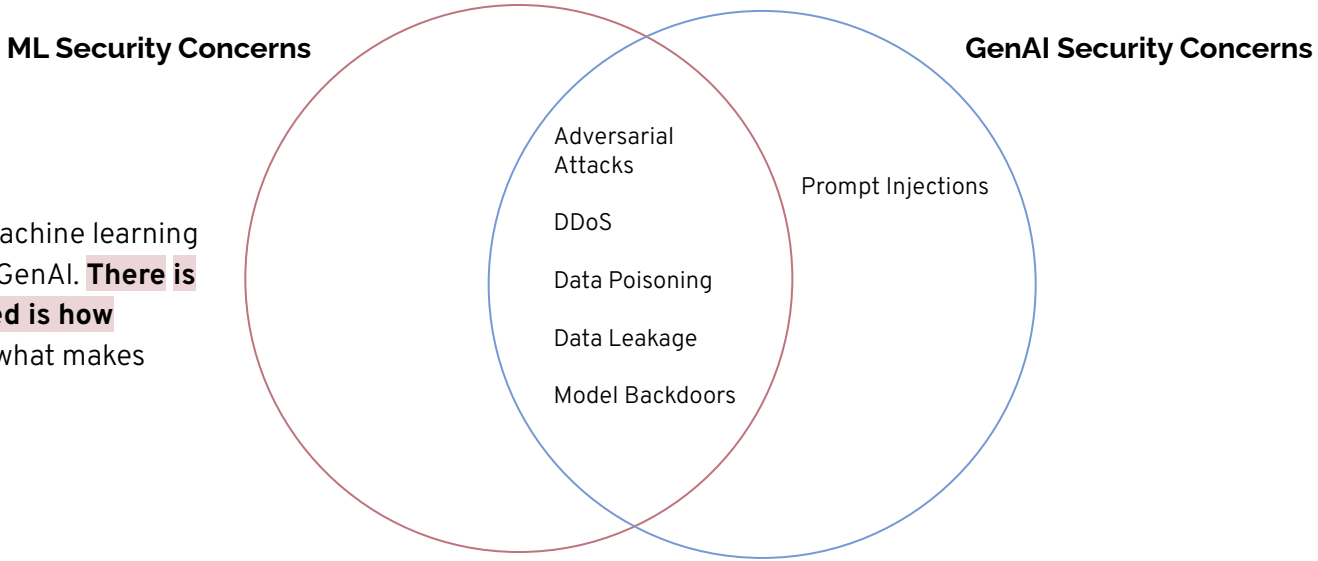
ALIGNED AI

LELAPA AI

AI Security: 2023 represented an AISec boom due to new fears around GenAI.

ML attacking ML becomes AI attacking AI.

Innovation in AI does not discriminate between defender and attacker. **As generative AI models become stronger, they also become more dangerous tools for malicious intent.** "ML to attack ML" (which has been around for a long time now) has already evolved into "AI to attack AI."



Every major concern from traditional machine learning returns in a slightly different form with GenAI. **There is very little new risk - what has changed is how abstracted the model itself is.** This is what makes security for GenAI models so difficult.

AI Security: the underlying trends for AISec invites more growth heading into 2024.

The path forward for AISec involves three main growth drivers.

The cybersecurity talent gap is wide and continues to widen.

- Difficulty to secure increases as AI systems become more abstracted, which is already happening as the barrier to deploy AI becomes lower and lower (e.g., low-code solutions, cloud giants offering ready-made pipelines, etc.).
- Firms need expertise in both modern-day AI and modern-day cybersecurity.

Investors (and the market) are unsure of what defines a strong AISec startup.

- The differentiation between the various AISec providers is extremely marginal, with all of them ensuring against the same set of risks. This means the rising tide of venture capital inflow is supporting the subindustry as a whole.
- 75% (six of eight) active AISec startups listed in EAIDB were able to raise funding in 2023.

Cybersecurity incidents involving GenAI models will continue to rise.

- While there hasn't been an extremely high-profile case of GenAI being compromised, academic institutions have highlighted their flaws several times.
- In 2023 alone, CMU researchers, Princeton researchers, and Adversa AI were just a few institutions that demonstrated successful, independent attacks on models from Anthropic, OpenAI, and Google.
- Enterprises worldwide will be watching for these incidents to determine their spending on AISec.
- Prior to 2023, there was a single paper on arXiv on "LLM attacks." **In 2023, there were 189 papers on the topic**.

Alternative ML: slow market growth but lots of unrealized potential.

The leading technological innovations in the AI space just don't have enough traction yet.

This could change in the near future with GenAI-specific product mixes.

Causal AI + LLMs

LLMs could learn to use causal principles to make reasonable estimates about the future. This is something they currently struggle with.

causaLens dynatrace

Neurosymbolic AI + LLMs

Neurosymbolic AI is a method of learning that is inherently interpretable and very high-performing. While too technical to dive into here, it is truly an exciting emerging technology.

UMNAI elemental cognition

Federated Learning + LLMs

For sensitive data applications, federated learning can provide a very efficient and private way of training LLMs on various silos. DynamoFL has their own 8B foundational model.

FedML DynamoFL

Market size estimates (millions, assuming no synergy)

Technology	2022/2023	2030	CAGR
Generative AI	29,000	668,000	48%
Causal AI	30	360	42%
Fed. Learning	119	298	13%

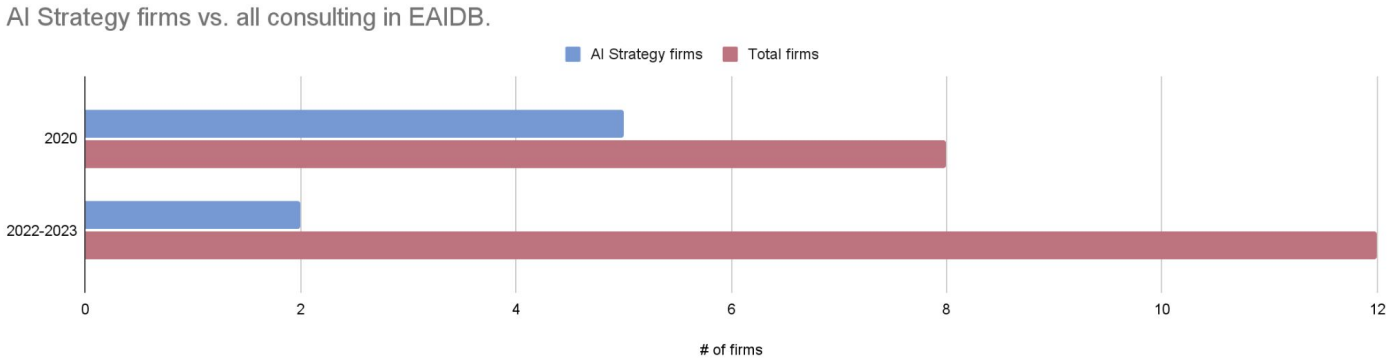
Market size data from Fortune Business Insights, Pioneering MarketDigits, and Grand View Research

Consulting: GenAI will require more than a plug-and-play approach.

Most consulting firms are still attempting to understand risk mitigation in a GenAI context.

While the technology behind GenAI is not completely different than traditional AI systems, enterprises and consulting firms require **re-education on what appropriate use cases are**. Most consulting firms in the RAI space are lean organizations that have had to scale their expertise on GenAI quickly over the last year.

In general, the number of "AI Strategy" consulting firms in EAIDB peaked in 2020 (when responsible machine learning was at its peak), but has since fallen. As firms scale their knowledge, they may begin to catch up in this regard.



A particularly interesting space is GenAI legal consulting, led by firms like INQ and Luminos. Another promising area is AI auditing with firms like BABL AI and Eticas, since the EU Act requires certification and validation.

INQ CONSULTING

LUMINOS.LAW

babl eticas CONSULTING

Open Source: the freemium model continues to dominate libraries and frameworks.

Upselling scale and performance once developers are hooked on an open-source framework is a proven business model.

This space is expected to heat up as GenAI development becomes more and more abstracted through open-source libraries. Some of the most successful names in GenAI-enablement have leveraged this business model.

Libraries enabling LLMs


LangChain


Hugging Face

vLLM


LlamaIndex

Libraries for validation, trust, etc.


trulens


Guardrails AI


Giskard


Confident AI

Much like AISec, **as deployment gets easier and more abstracted with names on the left, names on the right become more critical** to promote safety, transparency, and trustworthiness.

Thank you to our partners and the RAI community.

The RAI ecosystem is much larger than just startups - we celebrate each and every one of those working to enable responsible AI.

If you'd like to collaborate with us or submit your own company, please visit eaidb.org.

Appendix

Appendix: EAIDB's 2023 Highlights

EAIDB + 

EAIDB partners with an Australian research organization to profile responsible AI enablement in a quickly emerging market. This marks the second government entity that EAIDB has worked with, following on from Nordic Innovation in 2022.

BGV FORGEPOINT CAPITAL

EAIDB hosts a panel for Forgepoint Capital and BGV at the Securing AI Summit.

ASA FAccT 2023

EAIDB cited in FAccT '23 and the American Statistical Association.

EAIDB + EAIGG

EAIDB partners with EAIGG and several startup founders for a three-part miniseries.

Appendix: EAIDB's Value Proposition

Lead and Sales Sourcing

We've heard from several of our constituent companies that EAIDB's transparency and ability to filter, search, and compare companies within the same solution space has helped clients find products that match their needs.

Investment Sourcing

EAIDB draws attention from VC firms and founders alike. We have exercised our unique ability to make connections between the two parties for the sake of advancing the RAI space.

Marketing and Promotion

EAIDB has nearly 4,000 followers on LinkedIn and has received over 4,000 downloads on our various reports. We attract attention from the public, policymakers, founders, and investors alike.

Market Research

As a fully independent organization, EAIDB sits in a place of objectivity and methodical approaches. We do market research on behalf of governments and organizations to investigate and identify market opportunities, profile companies, and offer in-depth comparisons of technology used.