anch.AI®

Innovation humans can trust
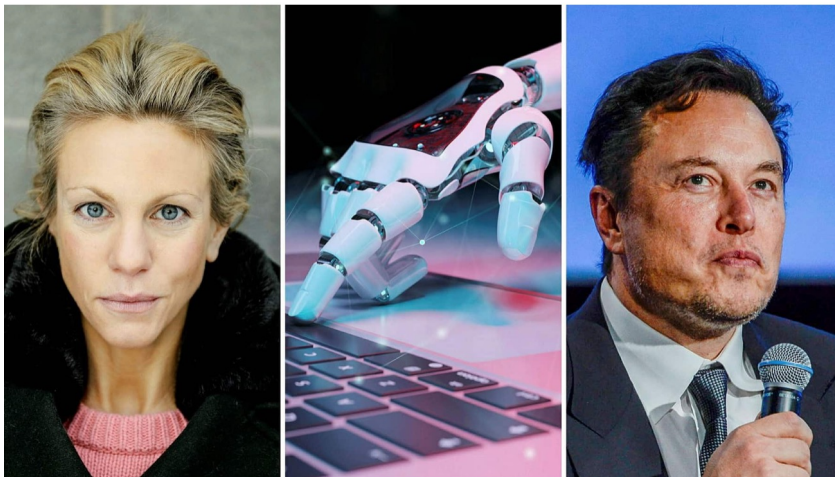
RELEASING THE REAL POWER OF AI

Too fast, too much, too soon?

anch.AI®

Svenska AI-experter sågar Musks varningsbrev
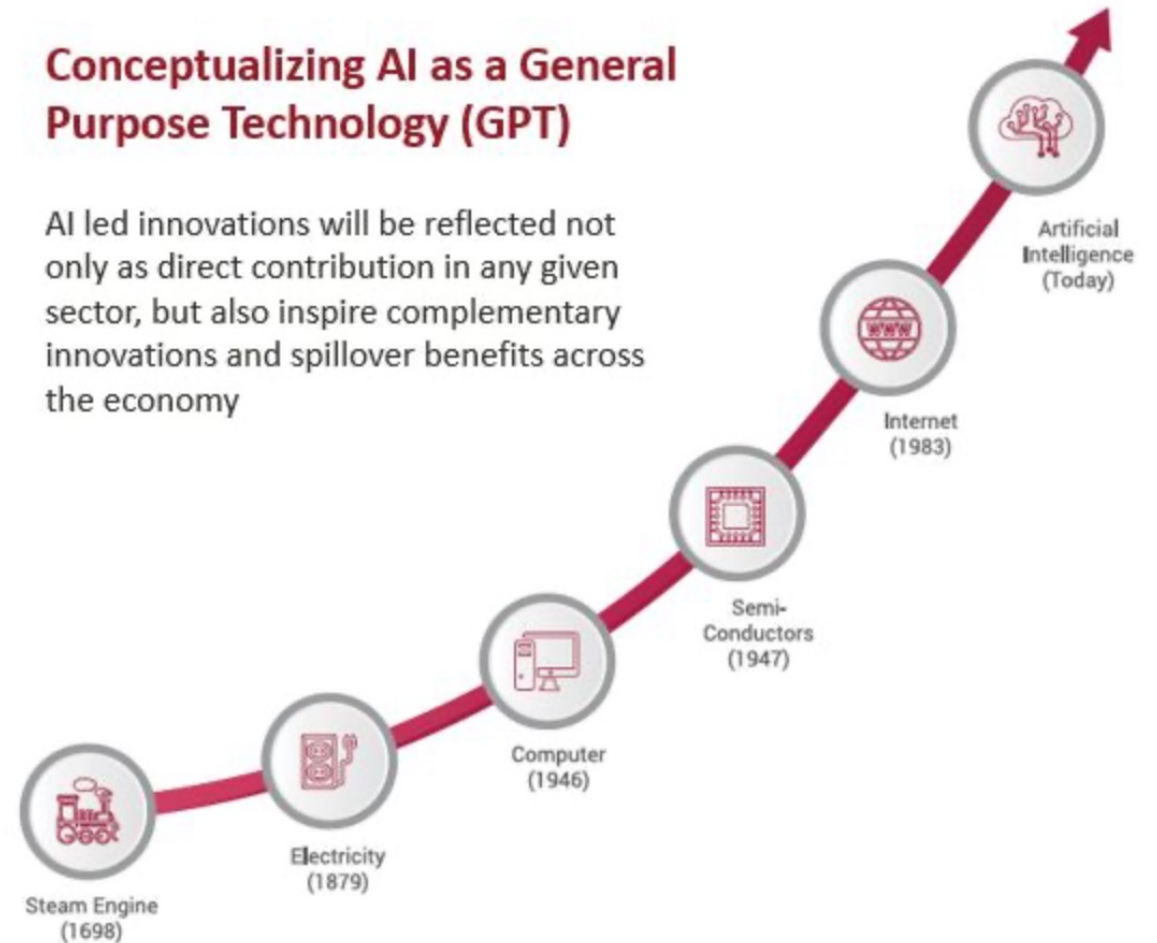
Lovisa Ternby • Publicerad: 30 mars 2023, 15:37

Anna Felländer tycker det största misstaget vore att sätta stopp för AI-utvecklingen. Foto: Jesper Frisk, Guillaume, Carina Johansen

Det pågår ett race "helt bortom kontroll" för att skapa en intelligens som "ingen – inte ens dess skapare" kan förstå, förutse eller kontrollera.

Undertecknarna av det öppna brevet som basunerades ut från Future of Life Institute igår skrädde inte orden när de varnade för att

Conceptualizing AI as a General Purpose Technology (GPT)

AI led innovations will be reflected not only as direct contribution in any given sector, but also inspire complementary innovations and spillover benefits across the economy

Artificial Intelligence (Today)

Internet (1983)

Semi-Conductors (1947)

Computer (1946)

Electricity (1879)

Steam Engine (1698)

Historical Instantiations of GPTs

anch.AI®

85% of AI projects are delivering misguided outcomes due to bias in data, algorithms or the teams responsible for managing them.

LOST AUTONOMY   **SOCIAL EXCLUSION**   **DISCRIMINATION**   **PRIVACY INTRUSION**

HARM TO SAFETY   **DES/MISS INFORMATION**

This poses ungoverned costly risks.

anch.AI

anch.AI®

| Values & principles | Accountability | Governance | Explainability | Transparency |

ORGANIZATIONAL FOCUS

USE CASE FOCUS

Establish to avoid:

Unintended *Pitfalls*

Establish to uphold:

Ethical AI *Declaration*

IMMATURE DATA & AI

Insufficient training of algorithms on data sets as well as lack of representative data could lead to incorrect and unethical recommendations.

BIAS OF THE CREATOR

Values and bias are intentionally or unintentionally programmed by the creator who may also lack knowledge/skills of how the solution could scale in a broader context.

MISUSE/OVERUSE OF DATA

The AI application/solution could be overly intrusive (using too broad or too deep open data) or it could be used for unintended purposes by others.

DATA BIAS

The data available is not an accurate reflection of reality or the preferred reality and may lead to incorrect and unethical recommendations.

01. Privacy

Everyone has the right to be in control of their private data.

02. Fairness

Everyone has the right to fair treatment.

03. Equality & inclusion

Everyone has the right to be treated equal, with equal access to resources.

04. Autonomy

Everyone has the right to personal freedom and autonomy.

05. Security

Everyone has the right to personal safety and security.

06. Information

Everyone has the right to be free from false or manipulative information.

Challenges governing risks around generative AI applications which do not present its training data or AI model

- Legal and Compliance Challenges
- Privacy and Security Risks
- Bias and Discrimination:
- Lack of Contextual Understanding
- Broader Ethical Concerns

anch.AI®

Heavy reporting and assessment requirements for "high risk AI" - for example in financial, transportation, public, education, recruitment and justice sector.

UNACCEPTABLE RISK

Social scoring, facial recognition, dark pattern AI, manipulation

HIGH RISK

Transportation systems, safety, employment, education access, border control, justice systems

LIMITED RISK

AI systems with specific transparency requirements such as chatbots, emotion recognition systems

MINIMAL RISK

AI enabled videogames, spam filters

STEP1

A high-risk AI system is developed.

STEP2

It needs to undergo the conformity assessment and comply with AI requirements.*

*For some systems a notified body is involved too.

STEP3

Registration of stand-alone AI systems in an EU database.

STEP4

A declaration of conformity needs to be signed and the AI system should bear the CE marking. **The system can be placed on the market.**

If substantial changes happen in the AI system's lifecycle

GO BACK TO STEP 2

anch.AI®

AI solutions are growing organically across organizations with NO consolidated or holistic view of ethical and legal risks. There is a lack of governance and consistence.
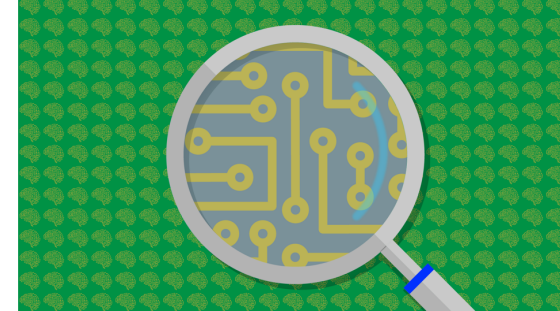
SCENARIO 1:

EXCEL AS GOVERNANCE TOOL – HIGH RISK EXPOSURE DUE TO LACK OF ITERATIVE APPROACH TO MANAGE PROGRESSION

SCENARIO 2:

AI GOVERNANCE TOOLS – IN A TECH SILO WITHOUT ABILITY TO REPORT TO AND ALIGN WITH NON-TECH DECISION MAKERS. FALSE SENSE OF SECURITY.

anch.AI®

Assess, audit and report AI risk from an integrated business, compliance and tech perspective

EMPOWER RELEVANT TEAMS WITH FULL CONTROL
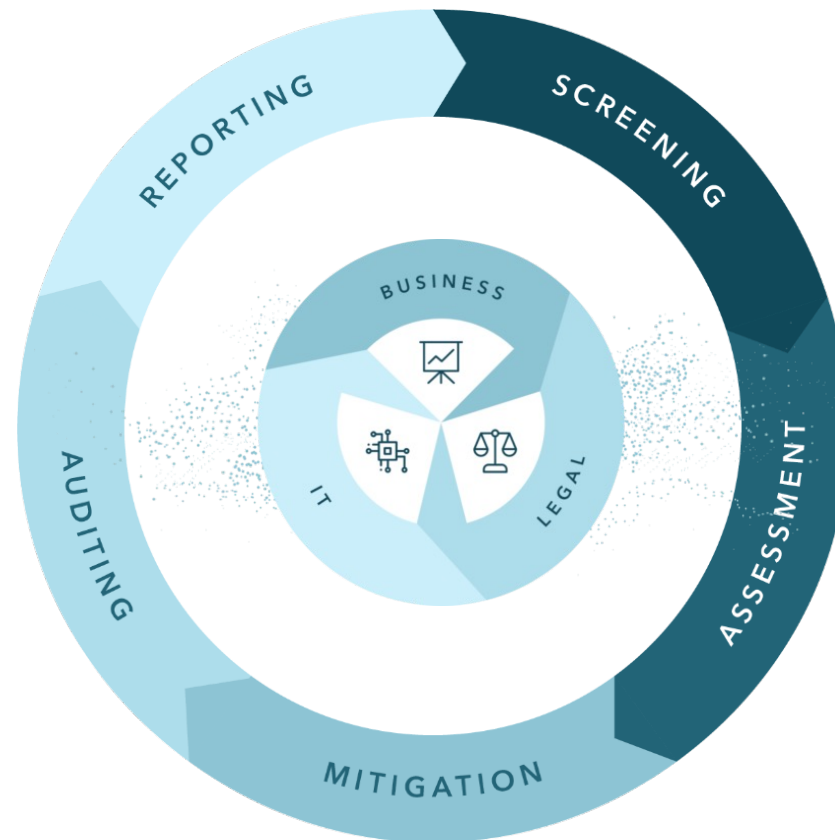
The benefit of using cross-functional governance – empowering all relevant teams

anch.AI®

For the tech team:
Speed and efficiency to accelerate AI implementation. Avoiding redesign of AI investments.

For the compliance team:
Full control of ethical and legal AI risks and avoiding EU AI Act sanctions up to 6% of revenue.

For the product team:
Cross-functional governance to reduce AI bias and get validation of responsible AI performance.
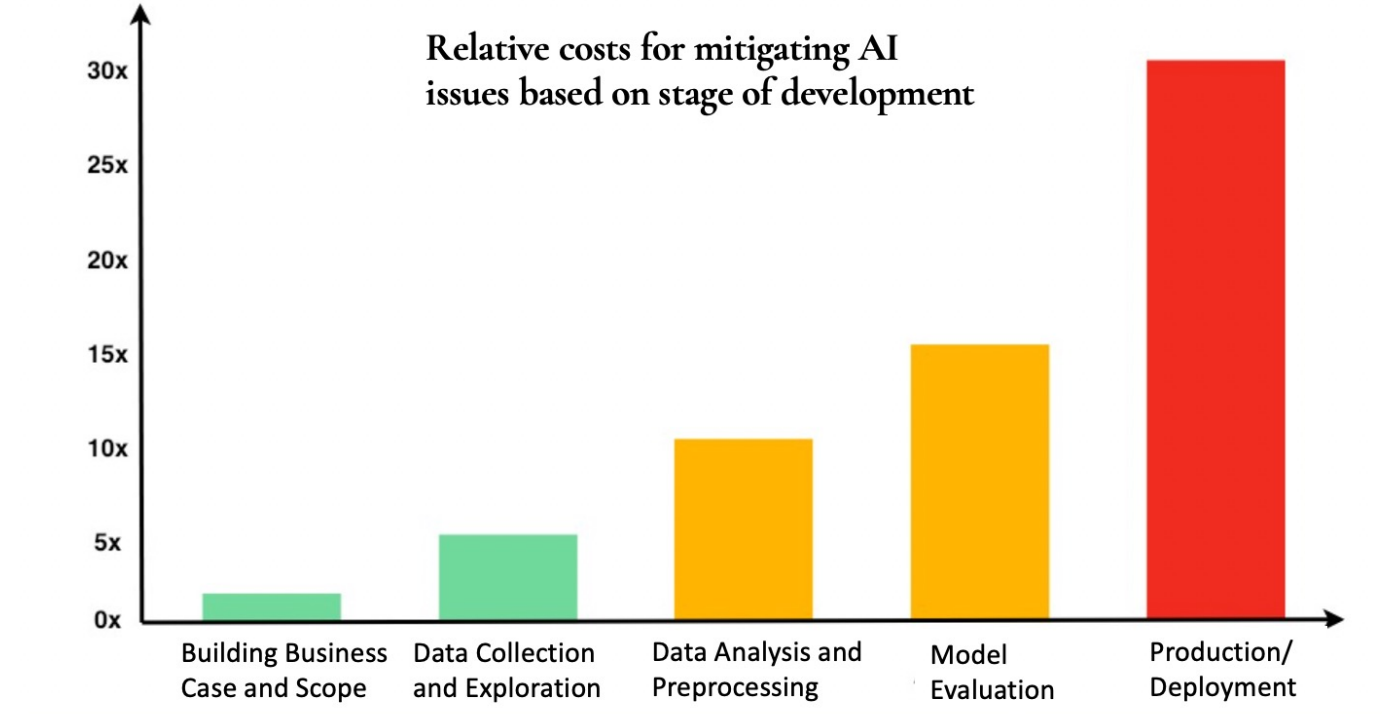
For the CEO:
Competitive edge through client and stakeholder trust. Avoiding reputational harm and lost customer trust.
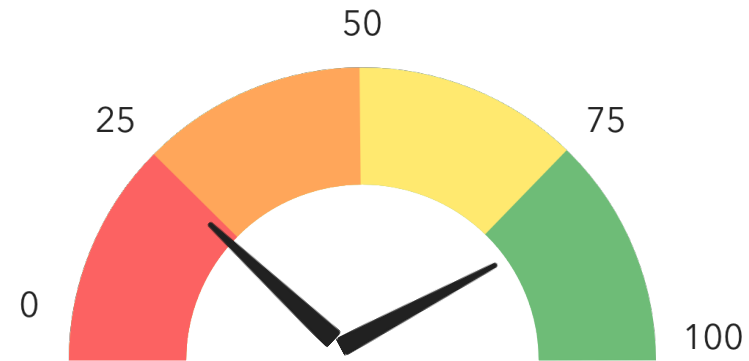
For the Board:
Reports in key risk indicators for ethical and legal AI risks.

Detect early, detect often, detect automatically. Go from *Unprepared* to *Reactive* to *Proactive* with AI Governance Maturity


Relative costs for mitigating AI issues based on stage of development

1. https://deepsource.io/blog/exponential-cost-of-fixing-bugs/, https://www.researchgate.net/figure/IBM-System-Science-Institute-Relative-Cost-of-Fixing-Defects_fig1_255965523

anch.AI®

OI.

DECLARATION ETHICAL AI

Privacy

Telecommunication corporation utilizing mobility data in product innovation

ETHICAL AI PERFORMANCE SCORE

anch.AI®

O2.

DECLARATION ETHICAL AI

Fairness

Pharmaceutical company developing vaccine decision support AI
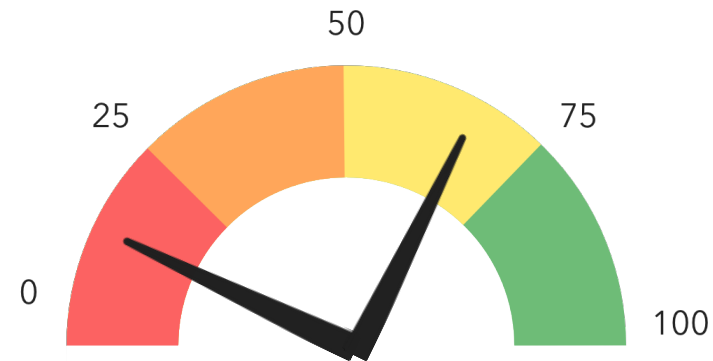
50

25

75

0

100
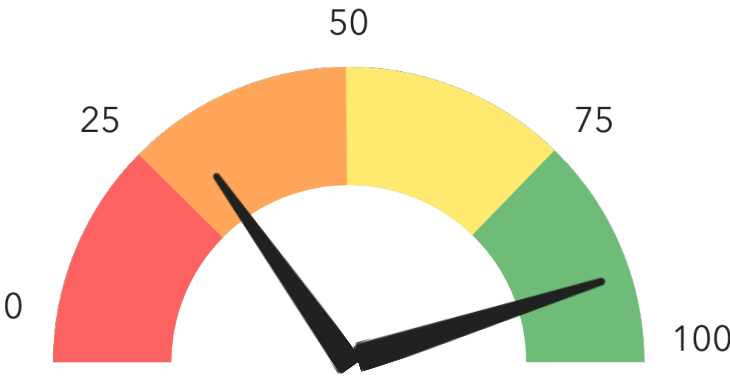
ETHICAL AI PERFORMANCE SCORE

anch.AI®

03.

DECLARATION ETHICAL AI

Equality and inclusion

Recruitment agency using AI to match job seekers with employers

ETHICAL AI PERFORMANCE SCORE

Thank you!

Contact us and visit our platform

info@anch.ai   anch.AI   twitter.com/anch_AI

anch.AI®

https://anch.ai

Mäster Samuelsgatan 36, 111 57 Stockholm, Sweden