EUROPEAN CENTRAL BANK

EUROSYSTEM

Trustworthy AI

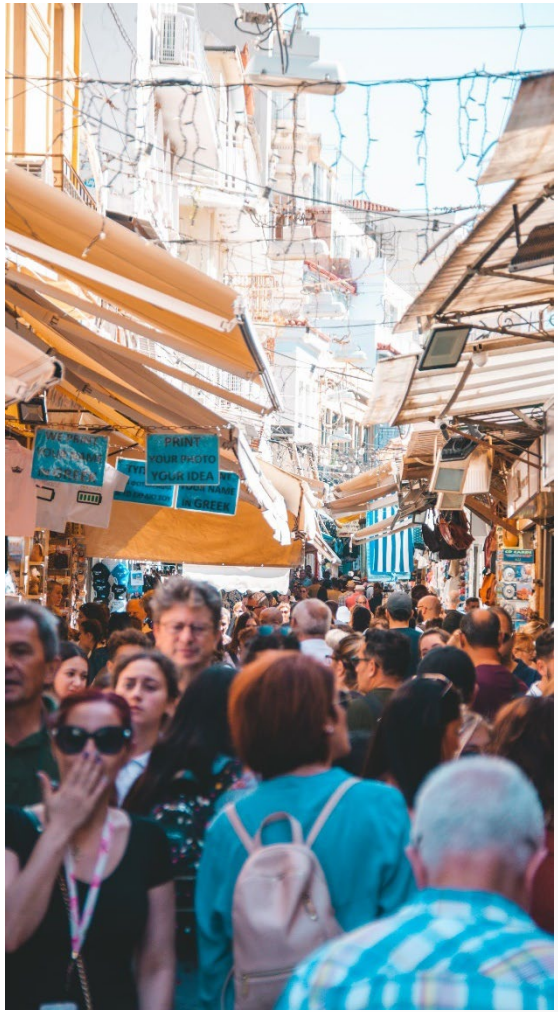


This image was created with the assistance of DALL·E 2

Maciej Brzezinski - Head of the ECB Data Office,
Jennifer Dooley and Pietro Graceffa

26 September 2023

The views expressed are those of the authors and do not necessarily represent the views of the ECB or the ESCB.



Maintaining Public Trust is crucial for Central Banks

Public trust is the **cornerstone** for fulfilling the ECB's primary mandate – **maintaining price stability**.

Recent studies show that a high degree of trust is crucial for anchoring the inflation expectations of private households near the level that the central bank defines as price stability.

Isabel Schnabel,
Member of the Executive Board of the ECB

https://www.ecb.europa.eu/press/key/date/2020/html/ecb.sp201216_1~9caf7588cd.en.html

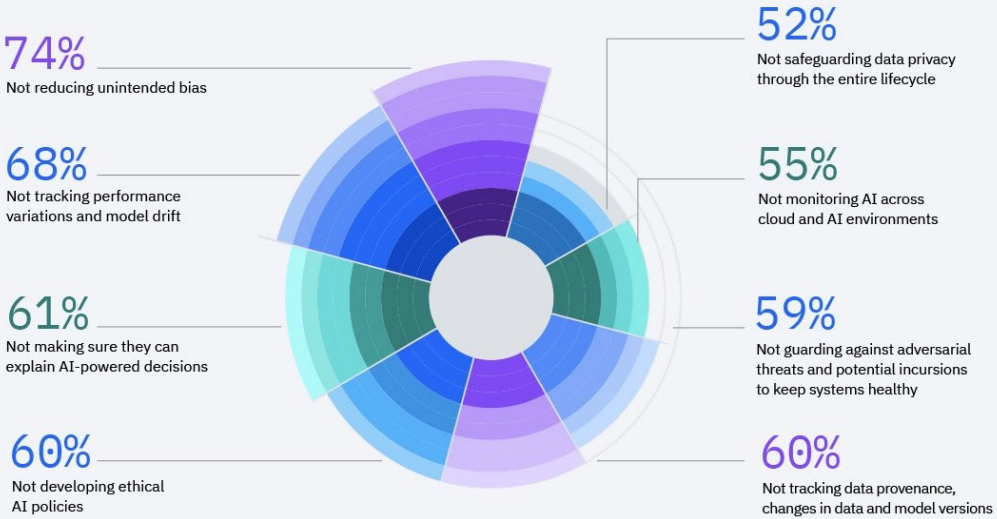
www.ecb.europa.eu ©

Adoption of AI

According to IBM Global AI Adoption Index 2022

AI adoption has been gradual as organisations work to implement holistic strategies, but the long-term trend is clear **only one in five companies today don't plan to use AI**

While organisations have embraced AI, **few** have made tangible investments in ways to **ensure trust or address bias**



New research commissioned by IBM in partnership with Morning Consult
https://www.ibm.com/watson/resources/ai-adoption

Does blocking Generative AI do any good?



The Daily

Sunday, August 30, 1993

Stop the Internet!

The internet is a dangerous new technology that threatens our privacy, our security, and our way of life. It is a tool for surveillance and control, and it is being used to spread misinformation and propaganda.

The internet is also addictive and isolating. It is keeping us from face-to-face interaction and from building real relationships. It is making us more passive and less engaged in the world around us.

I am calling on all people of conscience to stand up and say no to the internet. We must resist this technology before it is too late.

Here are some of the specific concerns that activists against the internet raised in the early 1990s:

The internet could be used to collect and store personal data about individuals, which could then be used to track them, target them with advertising, or even blackmail them. The internet could be used to spread misinformation and propaganda, which could undermine democracy and social cohesion. The internet could be used to facilitate criminal activity, such

- **Some organisations** have already **defined** the **way** how to **adopt AI**, they put in place **governance** and developed **policies** that are **formally approved** and **operationalised**
- **Some organisations** have **not started their AI journey yet** and they have not started defining their governance or guidelines for AI use
- **Majority of organisations are somewhere in the middle**, they are **considering** if they **should use AI**, they **evaluate risks and benefits**, experiment, **discus governance** structures, **preparing guidelines** and policies

It is not about rushing with AI adoption because everyone else is doing so

It requires careful considerations where to use it, should it be used, and how to ensure it is used in a trustworthy way

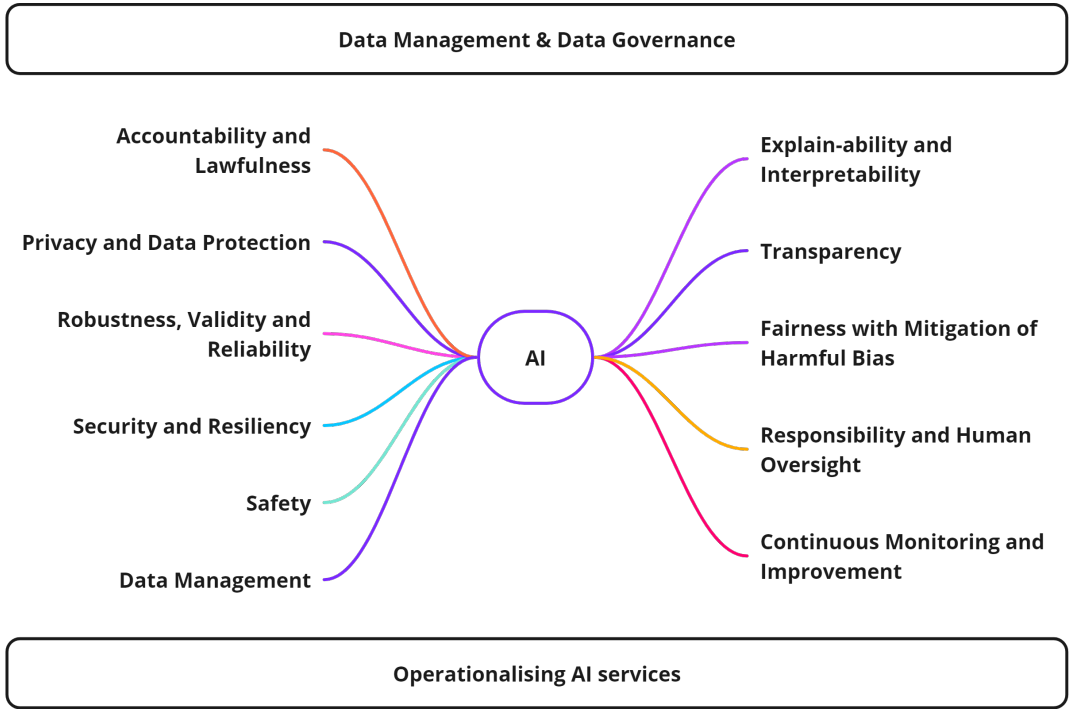
www.ecb.europa.eu ©

Principles of Trustworthy AI

- **Lawful, ensuring compliance with all applicable laws and regulations**
- **Ethical, demonstrating respect for, and ensure adherence to, ethical principles and values**
- **Robust, both from a technical and social perspective**, since, even with good intentions, AI systems can cause unintentional harm

Trustworthy AI concerns not only the trustworthiness of the AI system itself but also comprises the **trustworthiness of all processes and actors that are part of the AI system's life cycle**.

Challenges organisations are likely to face when adopting AI



Trustworthy AI is **essential** for **fostering public acceptance, ethical use**, and the **responsible deployment** of artificial intelligence in various domains, including **finance** and more.

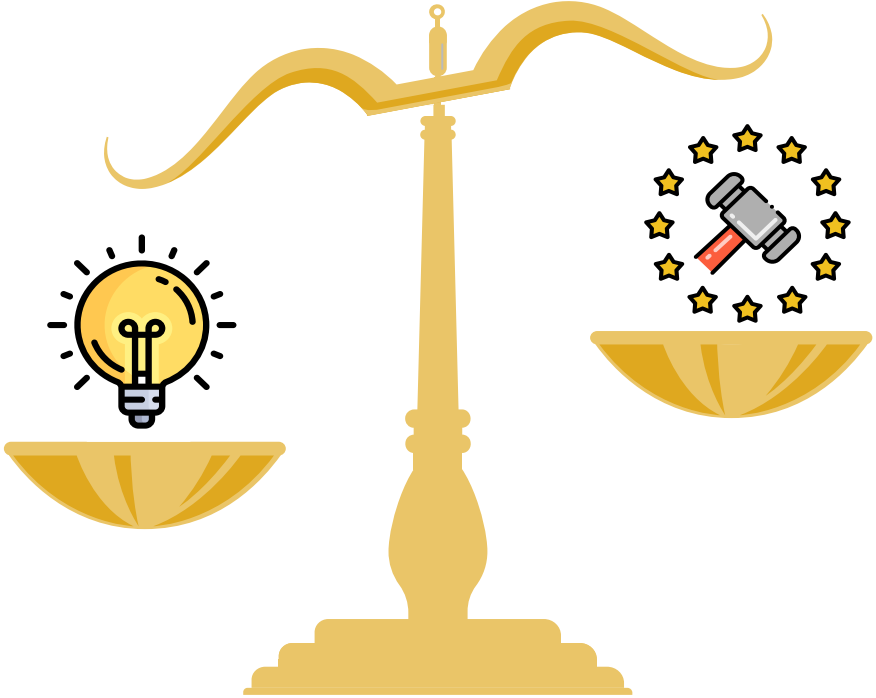
Organisations and regulatory bodies often establish guidelines and standards to promote trustworthy AI practices and ensure that AI technologies benefit society while minimising risks and negative consequences.

Chat-GPT 3.5

www.ecb.europa.eu ©

Challenges of AI Governance

- **Balancing regulations vs innovation**
- **Principles based vs detailed guidance**
- **Enhanced regulations may be needed to tackle bias and privacy issues**



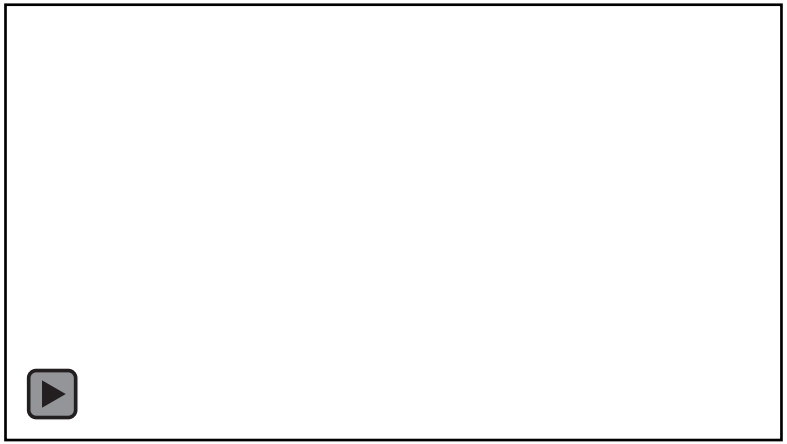
www.ecb.europa.eu ©

US AI Disclosure Act of 2023

Should it be disclosed when AI is used?

Generative artificial intelligence shall include on any output generated by such artificial intelligence the following: "**Disclaimer: this output has been generated by artificial intelligence.**"

Work is ongoing in the US Congress to adopt further regulations (e.g., the Blueprint for an AI Bill of Rights).

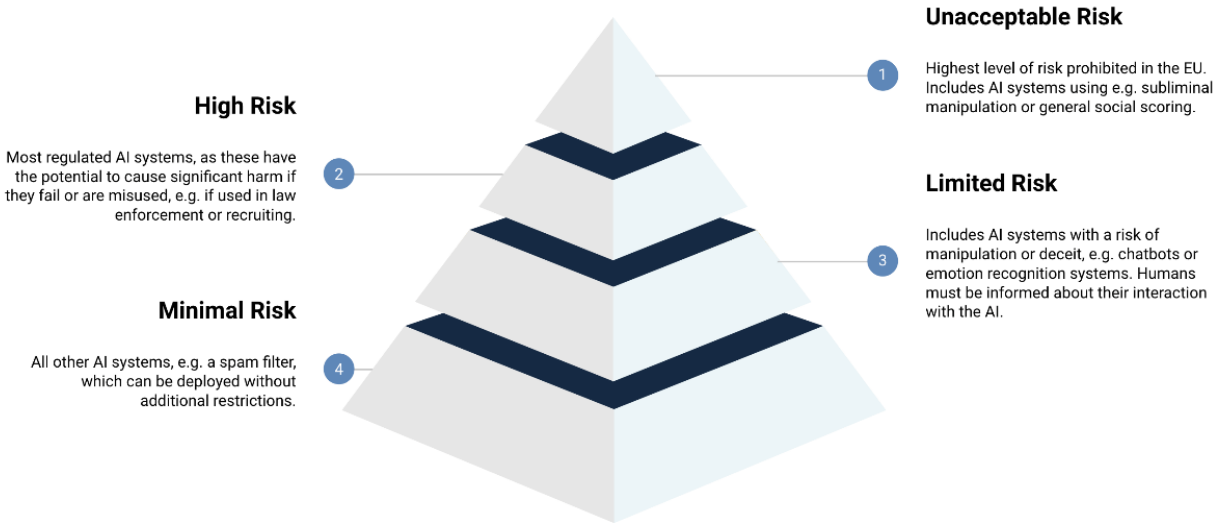


OpenAI CEO Sam Altman, Prof. Gray Marcus and IBM Vies President Ms Christina Montgomery testifies at Senate artificial intelligence hearing
https://www.youtube.com/live/TO0J2Yw7usM?feature=shared

EU AI Act

The AI **Act is a proposed European law** on artificial intelligence – a **risk based** comprehensive law on AI

Evet thought the proposed regulation has not enter into force yet, the current proposal provides insights into the future developments



Selection of Initiatives and Frameworks

UNESCO – First ever global standard on AI ethics adopted by all UN member states

OECD – Council Recommendation on Artificial Intelligence to benefits society as a whole

EU Ethics Guidelines for Trustworthy AI – Ethical guidelines with inclusion of human agency

EU AI Act – AI system risk assessment before establishing compliance requirements

US AI Disclosure Act 2023 – Disclaimer mandatory for outputs created by AI

US Blueprint for an AI Bill of Rights – provides an important framework for how government, technology companies, and citizens can work together to ensure more accountable AI

Canada AI and Data Act (AIDA) – this framework designed to encourage the responsible adoption of AI technologies by Canadians and Canadian businesses

UK A pro-innovation approach to AI regulation

New Zealand – Explicit guidelines for regulators from Privacy Commissioner & NZ Government

Monetary Authority of Singapore (MAS) released an open-source toolkit to enable the responsible use of AI in the financial industry

Latin America

Chile – the parliament has proposed a bill including AI legal and ethical concerns together with a national commission for AI

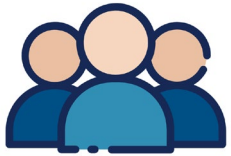
Peru – in June 2022 Peruvian congress approved a law to promote the use and development of AI. A national agency was created to supervise the development of AI using ethical and security principles

Colombia – in February 2023 government published recommendations for Trustworthy AI

Costa Rica – a bill on a regulation of AI

Argentina- in February 2023 published recommendations for reliable AI

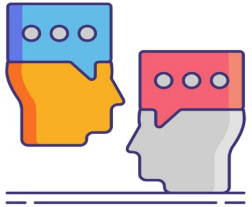
Elements of AI Governance for Central Banks



Wide range of stakeholders



Create mechanisms for stakeholders to provide input



Be transparent about plans for AI



Accountable for decisions



Literacy



Learn from others

www.ecb.europa.eu ©


thank you
danke
謝謝
ngiyabonga
شكراً جزيلاً
рахмат
teşekkür ederim
спасибо
gracias
dank je
tapadh leat
bedankt
dziękuję
obrigado
sukriya
kop khun krap
go raibh maith agat
merci
mochchakkeram
dakujem
terima kasih
grazie
arigatô
takk
hvala
mauuuru
sagolun
taiku
tack
barka
mersi
vinaka
blagodaram
matondo
paldies
grazzi
tapadh leat
хвала
asante
manana
obrigada
tenki
enkosi
djiere dieuf
tau
mamnun
chnorakaloutioun
gratias ago
gracies
sulpáy
akun
dankon
aciù
kiitos
dankie
faafetai lava
спасибі
Баярлалаа
nanni
nandri
dhanyavad
gracie
bayarlalaa
koszonom
sobodi
dèkuji
mesi
didi madloba
najis tuke
kam sah hamnida
rahmat
তোমাকে ধন্যবাদ
감사합니다
ありがとう
tanemirt
rahmet
xièxie
ευχαριστώ
diolch
dhanyavadagalu
shukriya
merce
мерси
trugarez
chokrane
murakoze
welalin
kia ora
seas
mahalo