



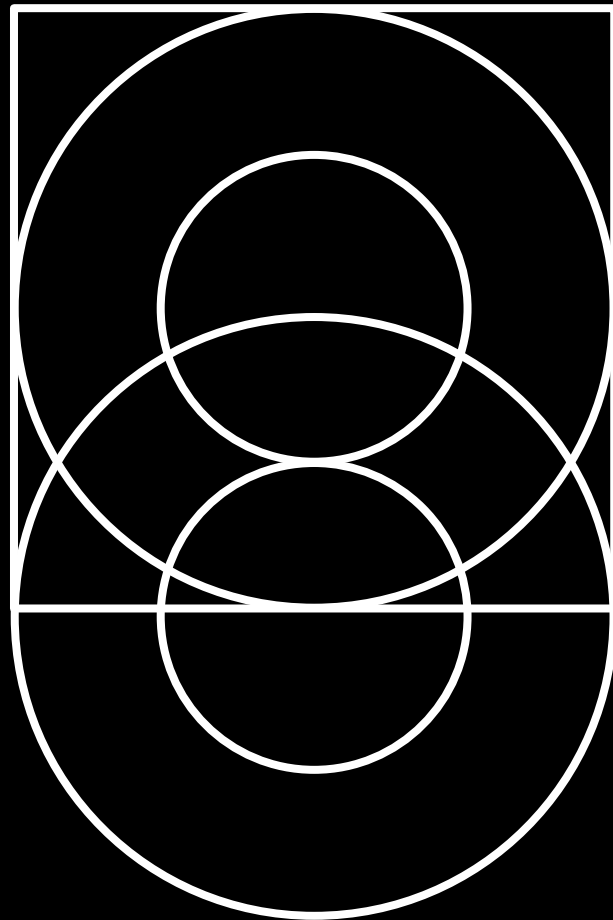
# How to ensure "Trustworthy AI" from a governance perspective

*Getting ready for the planned EU AI Act,  
also affecting Swiss entities*

**Andreas Blum**

Head of Digital & AI Consulting Practice at Unit8 SA

28.09.22



# Our mission as Unit8: Turning data into value

Unit8 is a leading Swiss data & analytics services company with a mission to help non-digital native companies transform data into value with a mix of **data science, analytics** and **AI**.

We are a “**one stop shop**” for solving data-related business problems using data & analytics, AI and custom software engineering.

## Focus industries



**Banking  
& Insurance**



**Manufacturing**



**Pharma  
& Chemical**



**Energy  
& Utilities**

Unit8.

## Services:

- Data & AI strategic advisory
- End to end custom solutioning
- Data platforms design & build up
- Data & ML Operations

## Key facts:

- Established in 2017
- 100+ employees and growing
- 5 locations: Zürich | Lausanne | Bern | Krakow | Wroclaw
- 120+ data & AI projects completed

## Selected Clients:

**MERCK**

**DAIMLER**

**Firmenich**

## Center of Excellence:

Gold  
Microsoft  
Partner



**Palantir**

## Selected Awards:



**INNOVAUD**

**digital**switzerland



**swiss made  
software**

# The widespread use of AI systems can lead to unintended, negative consequences and biases seen as unfair

2022

[Los Angeles Times](#)

California investigating racial bias in healthcare algorithms

2022

[The Swaddle](#)

How Facial Recognition AI Reinforces Discrimination Against Trans People

2021

[The Guardian](#)

Ex-Uber driver takes legal action over 'racist' face-recognition software

UK driver alleges his account was illegally deactivated when Uber software decided he was not who he said he was

2021

[Vox](#)

**AI's Islamophobia problem**

GPT-3 is a smart and poetic AI. It also says terrible things about Muslims.

2021

[Extreme tech](#)

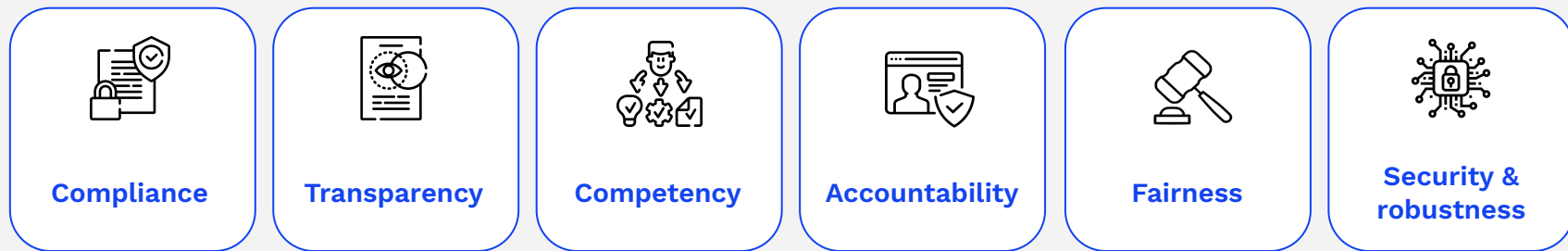
Predictive Policing Software Shown to Entrench Bias, not Address It

2019

[Independent](#)

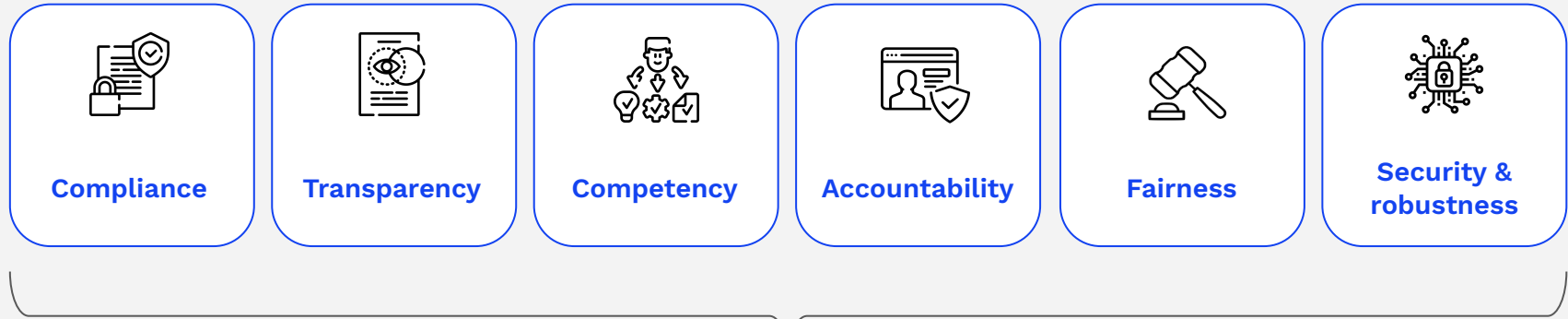
**Government 'deported 7,000 foreign students after falsely accusing them of cheating in English language tests'**

# To mitigate AI-related risks (and also drive business value), companies have started voluntary governance programs



**Pillars of AI governance which are relevant for building “Trustworthy AI”**

# To mitigate AI-related risks (and also drive business value), companies have started voluntary governance programs

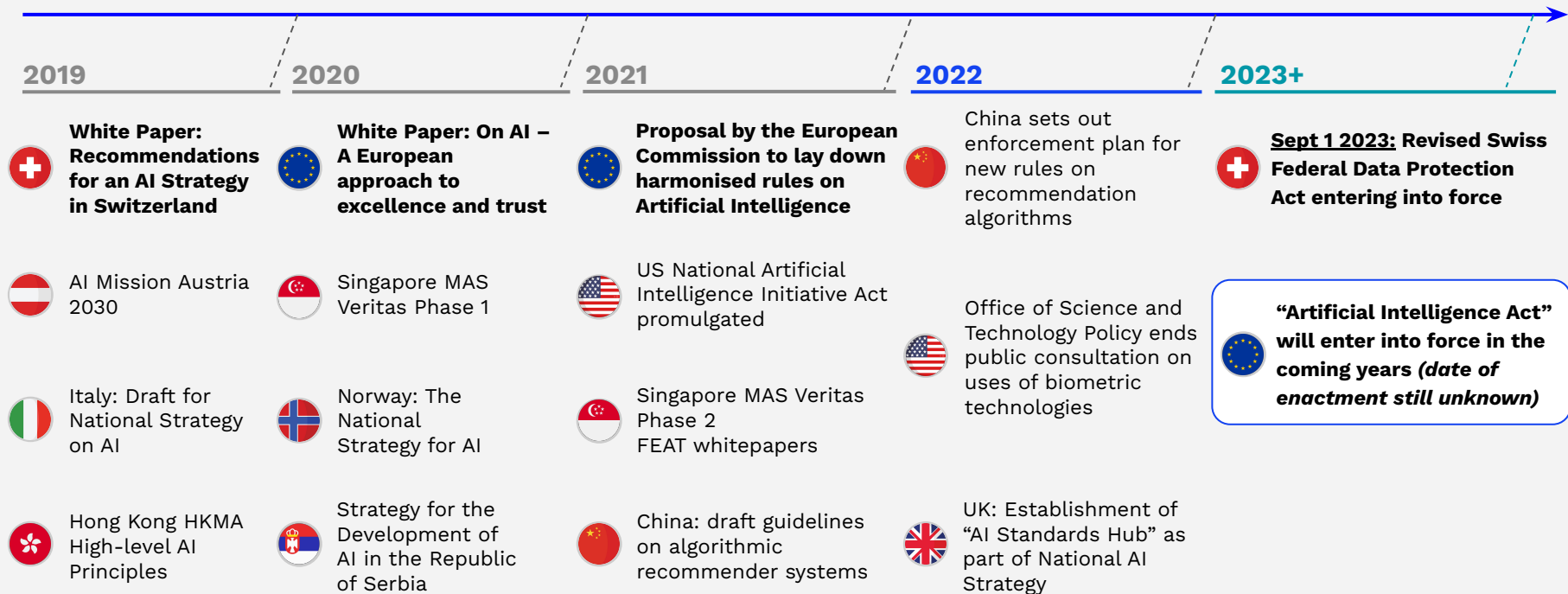


## Pillars of AI governance which are relevant for building “Trustworthy AI”

To better “manage” and govern AI, companies have started to ...

- build up **Centres of Excellence** or similar units dedicated to the management of **AI systems**
- develop **binding standards** and **policies** which govern the development and use of AI systems
- **register, classify** and **monitor** all **AI systems in use** by an organisation

# Whilst current AI governance is often still voluntary, the near future will see increasingly mandatory legal requirements

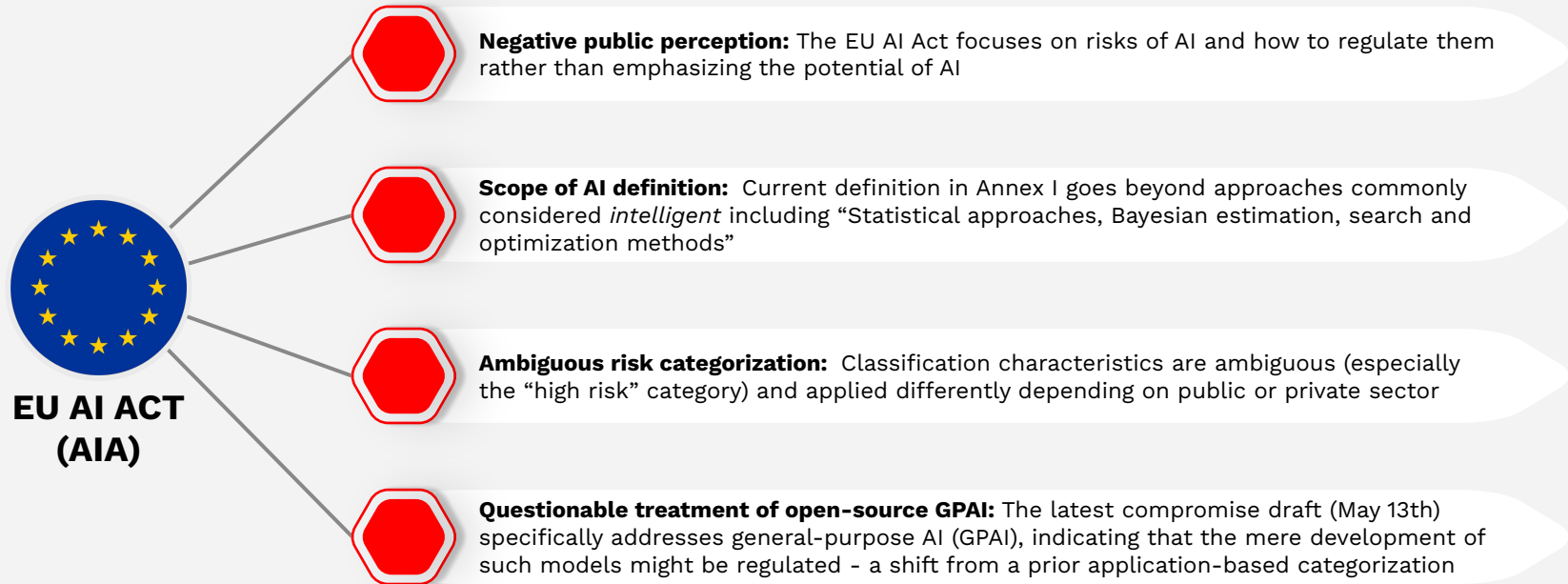


# The proposal by the European Commission foresees 4 risk categories of AI systems based on the application

Risk category	Example use cases	Legal requirements for AI system owner
1 Unacceptable risk	<ul style="list-style-type: none"><li>Government 'Social scoring'</li><li>Live identification based on biometrics</li></ul>	<ul style="list-style-type: none"><li><b>use of systems</b> which fall into this category is <b>prohibited</b> (very narrow exceptions apply)</li></ul>
2 High risk	<ul style="list-style-type: none"><li>CV-sorting</li><li>Credit scoring</li><li>Safety components of products</li></ul>	<ul style="list-style-type: none"><li><b>complex legal requirements</b> for the use of such systems apply (e.g. risk assessment, high data quality, activity logging, documentation, transparency, robustness)</li></ul>
3 Limited risk	<ul style="list-style-type: none"><li>AI Chatbots</li><li>Systems manipulating image, audio or video content</li></ul>	<ul style="list-style-type: none"><li><b>transparency obligations</b> apply to inform the user on the use of AI systems</li></ul>
4 Minimal risk	<ul style="list-style-type: none"><li>Spam filter</li><li>Inventory management systems</li></ul>	<ul style="list-style-type: none"><li>no additional legal requirements apply</li></ul>

- The regulation applies to providers of **AI systems put into service in the EU** or **AI systems whose output is used in the EU**
- Non-compliance can be **fined with up to 6% of global annual turnover** or **EUR 30m** (whichever is higher)

# The main criticisms concern ambiguous terminology and risk assessment, potentially hindering AI development





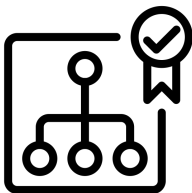
# Bringing it all together - How to prepare for the upcoming challenges?



Continue **monitoring** the **EU AI Act** legislative process



Perform **status quo assessment** to **identify current & planned AI systems** and the implied risk categorization



**Establish** a central **AI governance body** to **monitor & quality control** all **AI systems** in **use** and **mitigate** identified **risks**



**Develop** capabilities in-house and **build partnerships** for additional external expertise





Thank you - let's have  
a short discussion!

Zurich, Switzerland  
Tessinerplatz 7  
8002 Zürich

Lausanne, Switzerland  
Avenue d'Ouchy 4  
106 Lausanne

Bern, Switzerland  
Zentroom Bahnhofplatz 10B  
3011 Bern

Kraków, Poland  
Szpitalna 7  
33-332 Kraków

Wrocław, Poland  
Sucha 3  
50-086 Wrocław

Presenter contact details:  
[andreas.blum@unit8.com](mailto:andreas.blum@unit8.com)



# Selected sources for the interested reader

- Angler, A. (2022). *The EU's attempt to regulate open-source AI is counterproductive*. Brookings Techtank. Retrieved from <https://www.brookings.edu/blog/techtank/2022/08/24/the-eus-attempt-to-regulate-open-source-ai-is-counterproductive/>
- Brakel, M., Uuk, R., Prettnner, C. (Future of Life Institute). *The AI Act*. Retrieved from <https://artificialintelligenceact.eu/>
- Sioli, L. (2021). *A European Strategy for Artificial Intelligence*. CEPS webinar -European approach to the regulation of artificial intelligence. Retrieved from <https://www.ceps.eu/wp-content/uploads/2021/04/AI-Presentation-CEPS-Webinar-L.-Sioli-23.4.21.pdf>
- Scientific Foresight Unit (2022). *Regulatory divergences in the draft AI act*. European Parliamentary Research Service. Retrieved from [https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729507/EPRS\\_STU\(2022\)729507\\_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2022/729507/EPRS_STU(2022)729507_EN.pdf)
- Heikkilä, M. (2022). *A quick guide to the most important AI law you've never heard of*. MIT Technology Review. Retrieved from <https://www.technologyreview.com/2022/05/13/1052223/guide-ai-act-europe/>
- Meltzer, J., Tielemans, A. (2022). *The European Union AI Act. Next steps and issues for building international cooperation*. Global Economy and Development at Brookings. Retrieved from [https://www.brookings.edu/wp-content/uploads/2022/05/FCAI-Policy-Brief\\_Final\\_060122.pdf](https://www.brookings.edu/wp-content/uploads/2022/05/FCAI-Policy-Brief_Final_060122.pdf)