

Generative AI and language models: opportunities, challenges and policy implications

Les lundis de l'IA
Telecom Paris, 5 June 2023



Karine Perset



Today's Webinar on generative AI and language models



1. The OECD: A Hub for democratic market economies to cooperate
2. The OECD AI Principles, 2019
3. Overview of “generative AI”
4. Policy considerations of “generative AI”
5. OECD work on policy approaches, from regulation to standards to codes of conduct



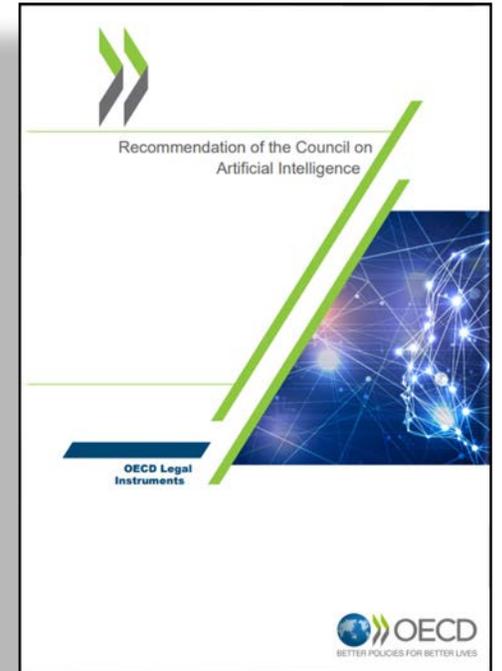
1. THE OECD: A UNIQUE HUB FOR INTERNATIONAL COOPERATION

- 38 democratic market economies
- To help rebuild Europe after WWII
- Intergovernmental organisation where governments & all stakeholders:
 - gather data, evidence, lessons learned
 - shape international policy standards
 - support implementation
- Allows multidisciplinary cooperation in all areas of policy except defense
- Sample OECD agreements: 1980 Privacy Guidelines; Polluter Pays; Anti-Bribery Convention; 15% minimum corporate tax



2. THE OECD AI PRINCIPLES, 2019

- Goal: create policy **ecosystem for trustworthy AI** that benefits people and planet
- **First** intergovernmental standard on AI, adopted 2019
OECD members, 8 partners
- **G20** AI Principles drawn from the OECD Principles



The OECD AI Principles



5 values-based principles for trustworthy, human-centric AI

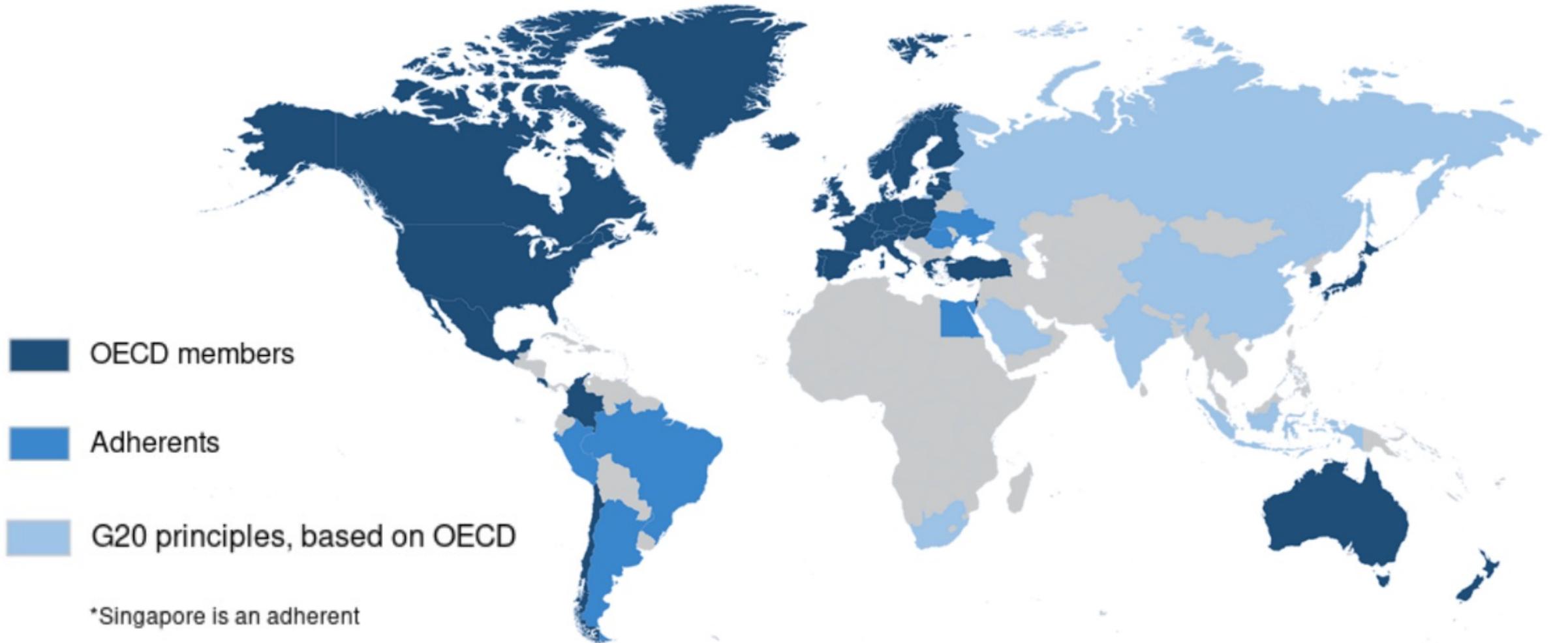
-  Benefit People & Planet
-  Human rights, values & fairness
-  Transparent & explainable
-  Robust, secure & safe
-  Accountable

5 principles for national policies, for AI ecosystems to benefit societies

-  AI research & development
-  Data, compute, technologies
-  Policy & regulatory environment
-  Jobs & skills, labour transitions
-  International cooperation & measurement

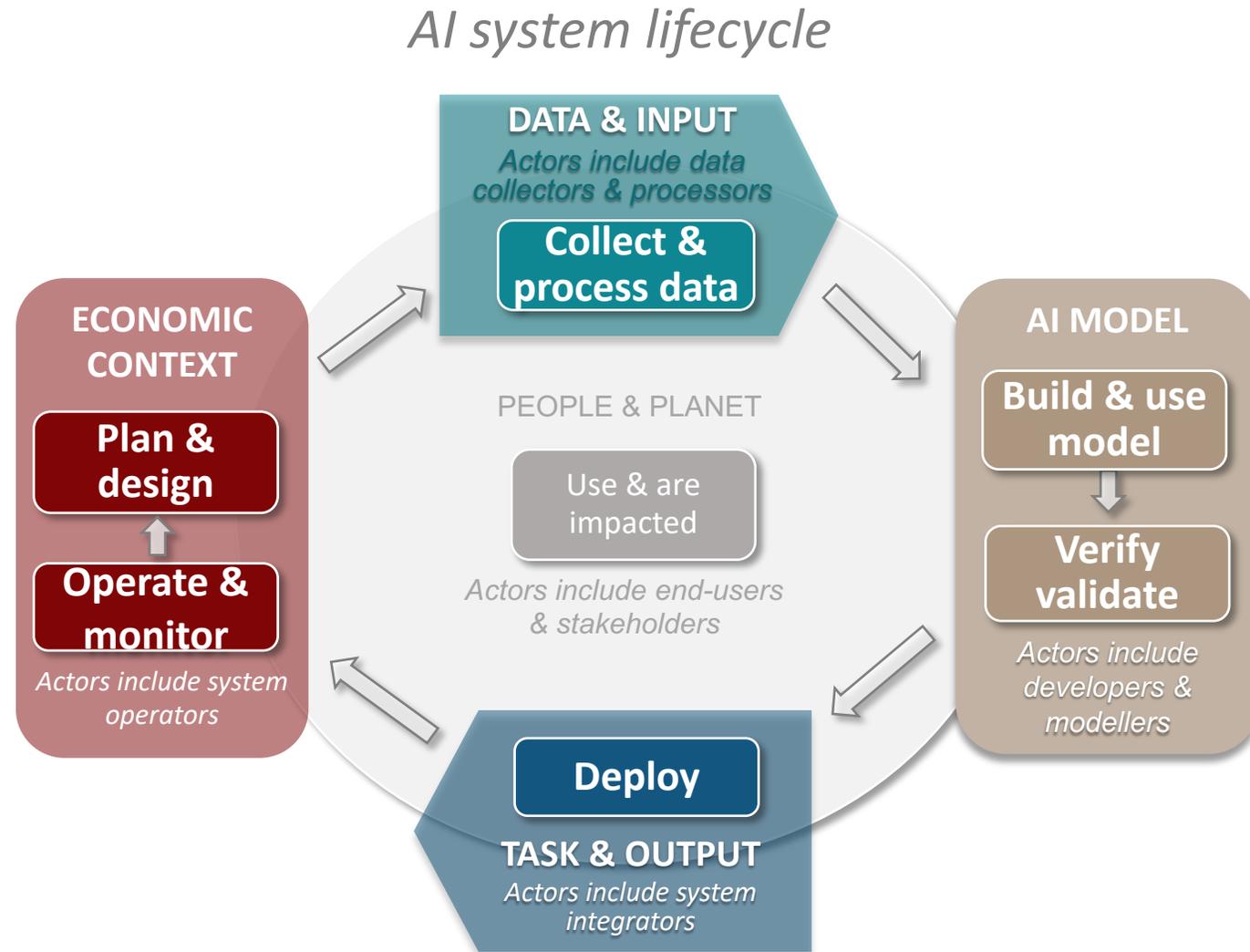


Governments that have committed to the OECD AI Principles



3. OVERVIEW OF “GENERATIVE AI”

- Can be understood as a form of AI that *produces NEW CONTENT as an output* (text, images, audio, video...)
- These systems are trained on massive amounts of data and use **machine learning**.
- Generative AI models receive input (data) and predict the next word or pixel to produce something new **with similar characteristics**.



What is or is not generative AI?

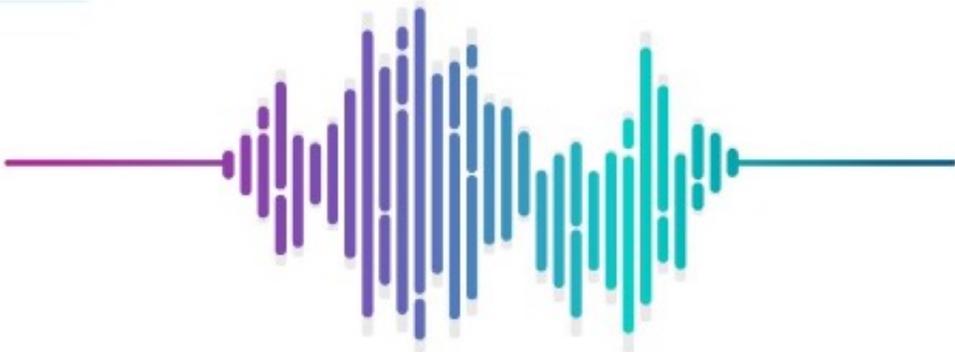
GENERATIVE AI



Artificial faces



Audio or voice synthesis



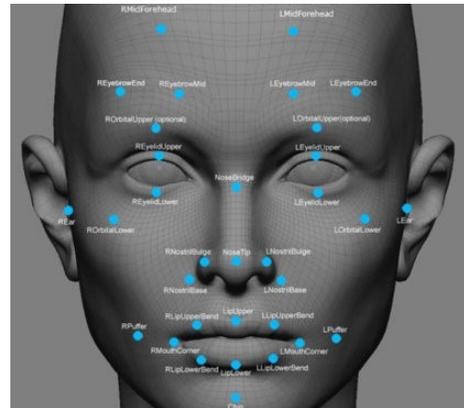
NOT GENERATIVE AI



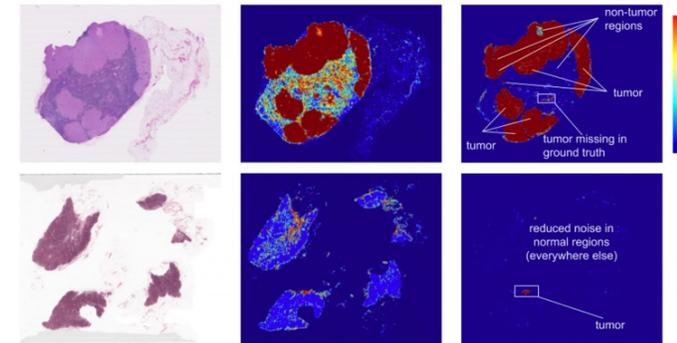
Driverless systems for A/Vs



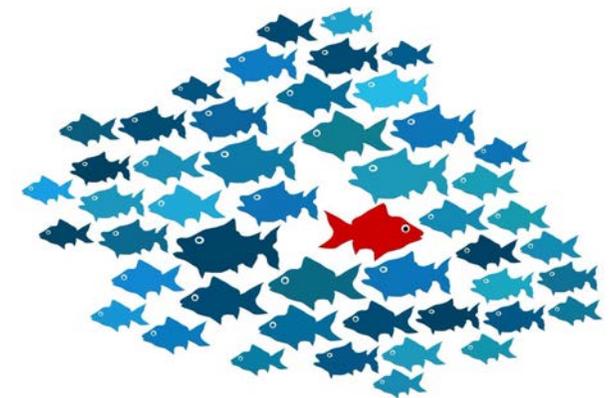
Recognition of objects, images, faces, music, voices



Cancer diagnosis tools



Anomaly detection



Generative AI handles content in many forms, from text to video to a combination of forms

Text-to-Image (T2I)	
Text-to-Video (T2V)	
Text-to-Audio (T2A)	
Text-to-Text (T2T)	
Text-to-Motion (T2M)	
Text-to-Code (T2C)	
Text-to-NFT (T2N)	
Text-to-3D (T2D)	
Audio-to-Text (A2T)	
Audio-to-Audio (A2A)	

There has been major progress in generative AI

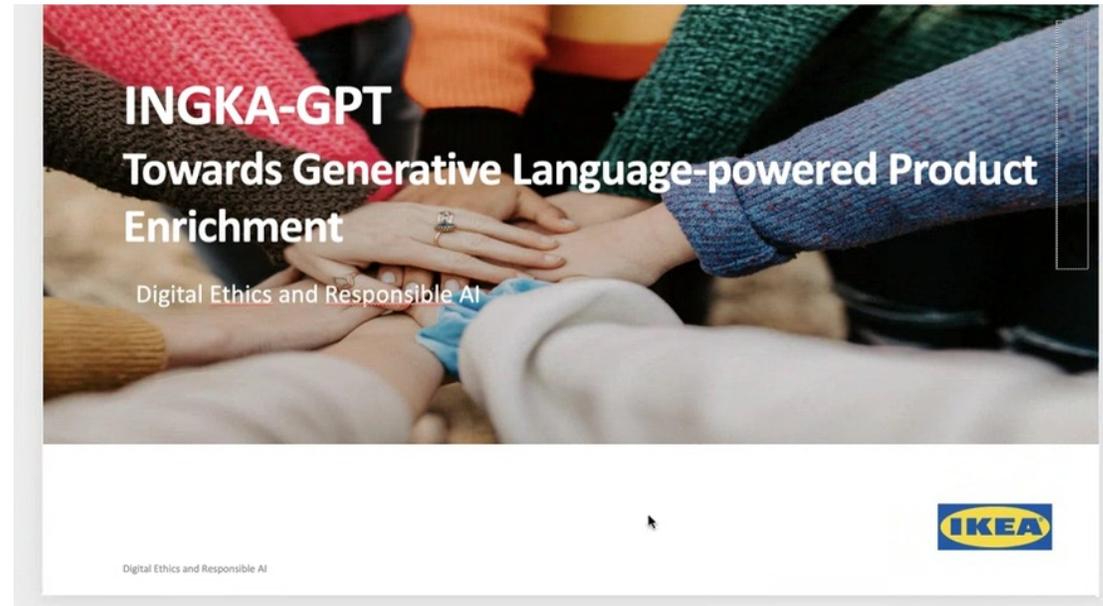
Prompt = “pixiv, hyper detailed, harajuku fashion.”



Midjourney
From
V1 2022
to
V5 2023



A GPT for everything?



4. Policy considerations: G7 leaders established the “Hiroshima process for generative AI

G7 Digital Ministers’ statement, April 2023 –

*“We recognise the need to **take stock** in the near term of the opportunities and challenges of (generative AI technologies) and to continue promoting **safety and trust** as these technologies develop.*

*We plan to convene future G7 discussions on generative AI which could include topics such as **governance**, how to safeguard **intellectual property rights including copyright**, promote **transparency**, address **disinformation**, including foreign information manipulation, and how to **responsibly utilise** these technologies.”*



SALIENT POLICY CONSIDERATIONS: OECD AI PRINCIPLES

Values-based OECD AI Principles

Benefit people & planet

- Availability of NLPs
- Environmental effects
- Financing concerns

Human rights, values, fairness

- Fairness and bias
- Privacy
- Dis-information, including foreign information manipulation

Transparent & explainable

- Specifications of black-box language models

Robust, secure & safe

- A double-edged sword for safety

Accountable

- Who should be held accountable and how?



Salient policy considerations: OECD AI Principles

Values-based OECD AI Principles

Benefit People & Planet

- Developing solutions that effectively mitigate NLP risks while fostering beneficial adoption.

Data, compute, technologies

- Providing data, compute and technology knowledge.
- Intellectual property rights of training data & whether/who has IP for AI-generated content.

Policy & regulatory environment

- Development of national action plans, research centers and/or language resources.
- Governance of generative AI

Jobs & skills, labour transitions

- Adapting to labour market upheaval
- Equipping people with the right skills

International cooperation & measurement

- Who should be held accountable and how?



A few OECD initiatives to advance AI governance for genAI

- Tracking developments: investment; R&D; jobs; education etc. [OECD.AI/trends-data](https://oecd.ai/trends-data)
- **Risk assessment** tools and guidance: [OECD.AI/tools](https://oecd.ai/tools).
- International **foresight and anticipatory governance**, including for networked AI systems.
- Focus on smart regulation via **regulatory experimentation**
<https://oecd.ai/en/work/sandboxes>.
- Investigating mechanisms to ensure the responsible use of AI, such as **certification**.
- Developing global AI incident monitor to **track AI incidents** and prevent harms from re-occurring.



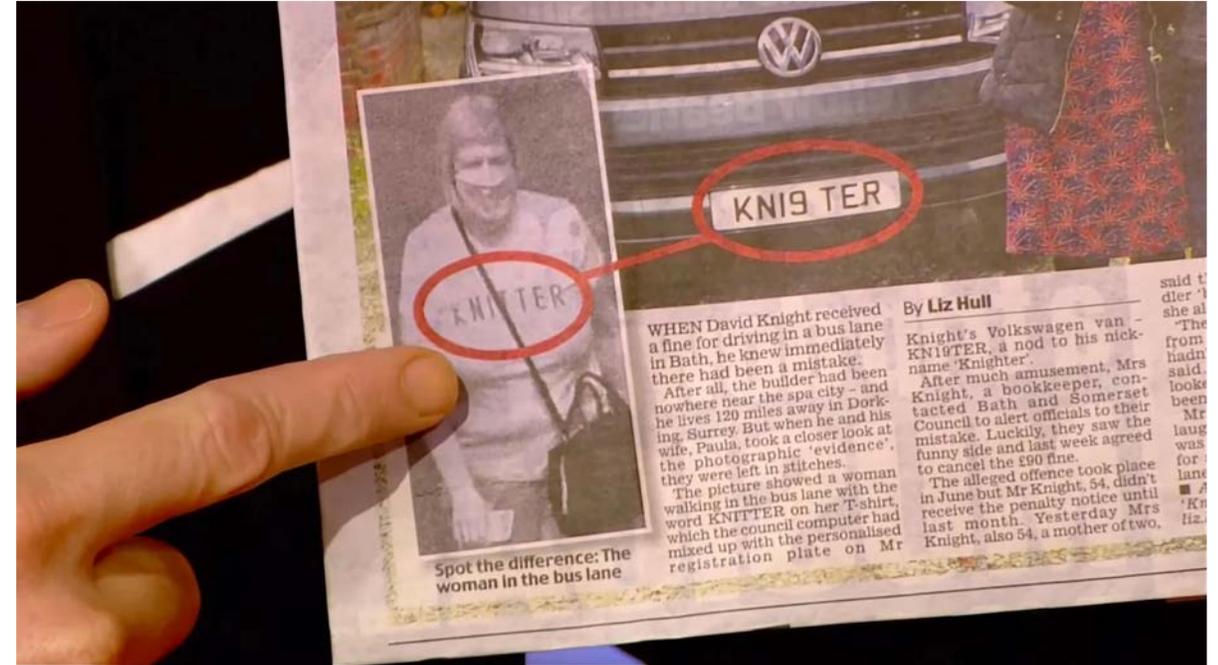
AI incidents come in many shapes!

Picture on bus mistaken for human



AIID Incident 36: Picture of woman on bus billboard led to woman shamed for jaywalking and credit score dropped in China

Human mistaken for car



AIID Incident 171: Traffic camera read text on pedestrian T-shirt and interpreted it as a license plate and fined the license plate owner in the U.K.



6. Generative AI: paths being explored

Soft law & self-regulatory guidance

US, EU, UK, JP:

- * Technical tools
- * Codes of Conduct
- * Model cards

Standards & risk frameworks

US:

- Public Assessments to evaluate model alignment w/ AI Bill of Rights Blueprint
- NIST AI Risk Management Framework
- NTIA RFC on AI accountability mechanisms

Use, interpretation & application of existing rules and regulators

UK:

Leverage existing regulators eg HSE, HRC, CMA

EU: European Data Protection Board's Task Force on ChatGPT to promote collaboration and share information

AI-specific regulation / hard law

EU: Proposed EU AI ACT (Parliament proposal):

- Transparency requirements
- Safeguards
- Summary of training data protected under copyright law

CA: Artificial Intelligence and Data Act (*AIDA*)



Selected policy initiatives to promote AI with national language

Digital language resources for smaller or indigenous languages

- **Canadian** indigenous languages technology project
- **France** Pour des IA Francophones “PIAF” -- building French-language datasets for AI
- **Japan** Global Communication Plan 2025
- **UK** Welsh Language Technology Action Plan
- **German** Federal Ministry for Economic Affairs and Climate Action created the SPEAKER project to develop a “made-in-Germany” voice assistant platform

NLP research centres and collaborative platforms

- **German** Society for Computational Linguistics and Language Technology

Cross-border knowhow/best practice sharing and facilitating interoperability

- **France** Cité internationale de la langue française et de la Francophonie and International Centre for Digital and language technologies





MERCI !

FOLLOW US:
@OECD-AI

