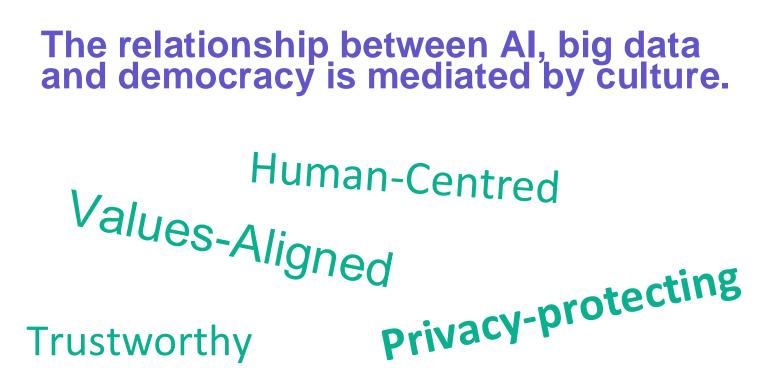**A culture-centred approach to AI "trustworthiness" in the context of democratic exchange**

Jennifer Edmond
Trinity College Dublin, KT4D Coordinator

Funded by
the European Union

The relationship between AI, big data and democracy is mediated by culture.

Human-Centred

Values-Aligned

Privacy-protecting

Trustworthy

Concepts of trust, trustworthiness and responsibility are **culturally determined**, *as is our understanding of the limits of personal freedoms, and the line between the private and public spheres. ... a narrow view of the role of culture can misrepresent its pervasive quality, however: hiring policies [alone cannot] make software align with human-centric values. The cultural positioning of the users of a technology, and the manner in which this shapes their perceptions and actions, are not taken into consideration [in the AI Act].*

Influence Factors

Culture: distributed control; flexible and adaptable over time; perceived as part of identity	Control: Seeking to maximise benefit without bias	Technology Development: centralised control and knowledge; fast-changing rather than adaptable; perceived as a consumer choice

Democracy and Civic Participation

POLICY BRIEF

Culture's Role in Navigating Technological Change:
The KT4D perspective on recent developments in European AI policy

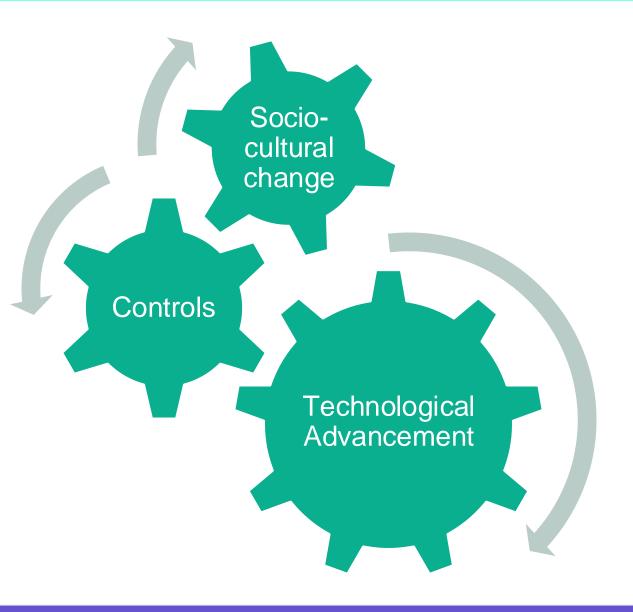DOI Number: 10.5281/zenodo.10725945

FEBRUARY 2024

Culture also frustrates some of technology development's vested interests, however. Evgeny Morozov has warned us of the dangers of 'technosolutionism' by which we seek to solve problems wrought by technology with more technology.
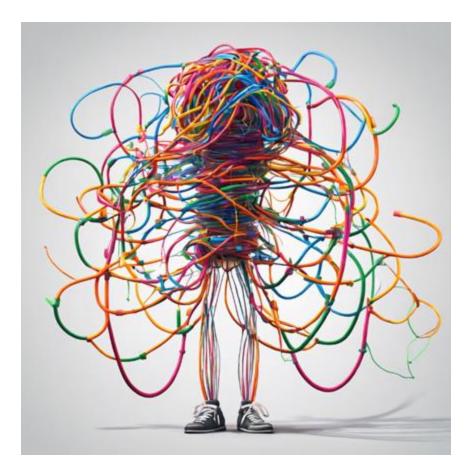
Three intersecting perspectives

- Technological affordance, but also business models and economic drivers, and in particular Human/ System Interactions

- The ability of culture to adapt, but risk it can be hacked

- Controls as available at multiple levels

From Human(s) in the Loop to Democracy in the Loop?

Can a technology really be considered 'trustworthy'

"Trust is an ability to rely on somebody to do what they have said they will do, even when no one is watching them" – Simon Longstaff AO

To extend trust, we need to feel we understand an alignment of values and capacities.

In opaque, complex knowledge processing systems, neither of these conditions can reasonably be met.

We should not confuse confidence in system function with the all-important social risk-taking of interpersonal trust.

The Digital Democracy Lab as an Experience, not a Platform

- Stemming from a skepticism regarding current 'human-in-the- loop' approaches (including participatory ML and anything 'by design.')

- Based upon the concept of 'meaningful friction' as a democratic and a design principle

- Three 'loops': input data, resulting profiles, policy language (AI approaches: RAG and LLMs).

- System is inherently, intentionally, delivered as fragments tied together by a strong facilitation guide.

- Slightly different interaction models for each of our three different user communities (policymakers, citizens and technology developers) taking place through October-November 2024.

Thank you for listening!