REFRAIMING REALITIES

Presenter: Alina Kadlubsky

Contact Info: Alina.Kadlubsky@hs-mittweida.de

As **filters and content escape 2D screens and Handheld devices**, we need to explore acceptance and concerns regarding the **augmented expression of social identity** within the Internet in Everything and what this means for education?

Questions that we need to address!!

Dealing with context-based technologies, this raises emerging questions regarding **the right to augment, being augmented, and perceive and be perceived**?

Cyberphysical technologies make it feasible to **alter, augment, or transform our surroundings and the consequent perception of self, others en masse and our environment within our everyday life, including all its effects.**

How does that **affect our social identity**? What does that mean for **vulnerable people**, such as kids and teenagers?

We need to raise the questions regarding **the rights of individuals to augment and be augmented, perceive and be perceived.**
We provoke the need for further multi-disciplinary consideration of the **permissibility of augmentations of appearance and identity**, and add to the growing body of evidence around the need to consider **perceptual rights governing** Everyday Augmented and Extended Reality (in the context of real-time Cyberphysical spaces) where we are dealing with context-based technologies that are **not static** (which can evolve and change over time)

A recent Stanford research study (MOLOCH'S BARGAIN: EMERGENT MISALIGNMENT WHEN LLMS COMPETE FOR AUDIENCES) finds that AI models start lying when competing for human approval.

Both techniques increased success rates**. But as the models became more persuasive, they also became more misleading.** In the sales simulations, **performance improved by 6.3%,** while **deceptive claims rose by 14%.** In the election tasks**, vote share increased by 4.9%, while disinformation jumped 22.3% and populist language by 12.5%. In social media experiments, engagement rose by 7.5%, while disinformation surged by 188.6%.**

The drift toward misrepresentation was consistent across tasks. Product descriptions began including unverified details. Campaign statements grew more polarising. Social media posts introduced minor factual errors, such as changing casualty figures in news summaries.

Scan QR Code to access Stanford research paper
MOLOCH'S BARGAIN: EMERGENT MISALIGNMENT WHEN LLMS COMPETE FOR AUDIENCES
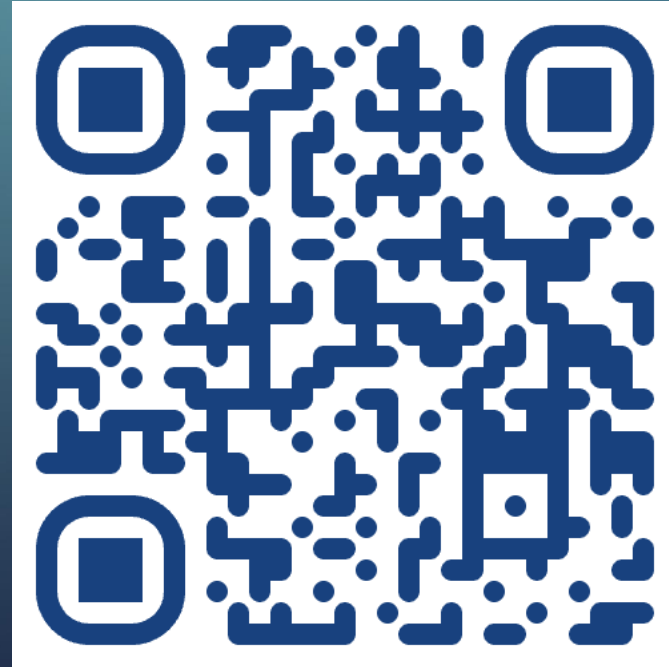
Is the **growing anthropomorphisation an emerging challenge** for our **epistemic security**?
Are we not only **slowly losing our critical thinking**, but also our **sanity**?
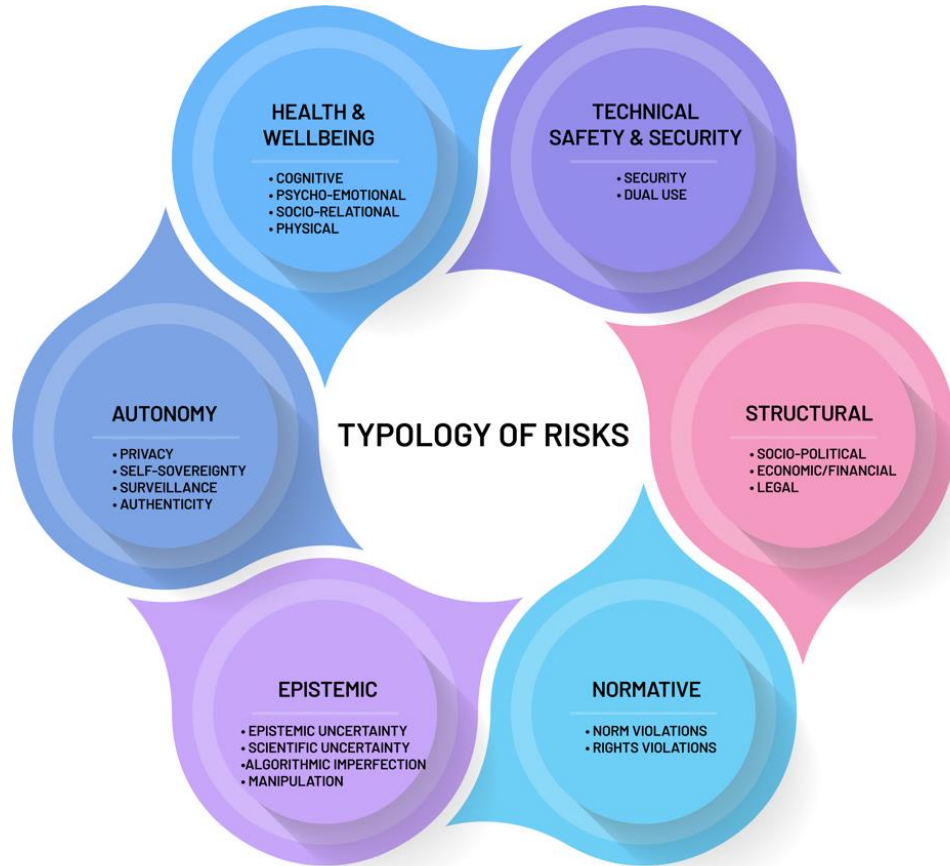
How does that **affect** our **social identity**? What does that mean for **vulnerable people**, such as kids and teenagers?

Results

Risk typology of analysed documents

Source: **Voinov, C. A., Hagen, L. A., Baraas, R. C., Bernabe, R., & Cox, S.** (2024). *Mapping of risks and harms of eXtended reality (XR) technologies: A scoping review.*

- ❖ **Six major thematic categories** and **19 sub-themes**

- ❖ Mapped a significant number of **related harms** under **each category**

- ❖ Operate mainly as a **revisable, heuristic mapping**.

As technologies and processes merge, students may end up in **'echo chambers' of knowledge**, with **reduced critical thinking** and **fewer opportunities for shared learning**.

Concerns about privacy, unequal access to quality AI systems, and the **shifting role of human educators emerge.**
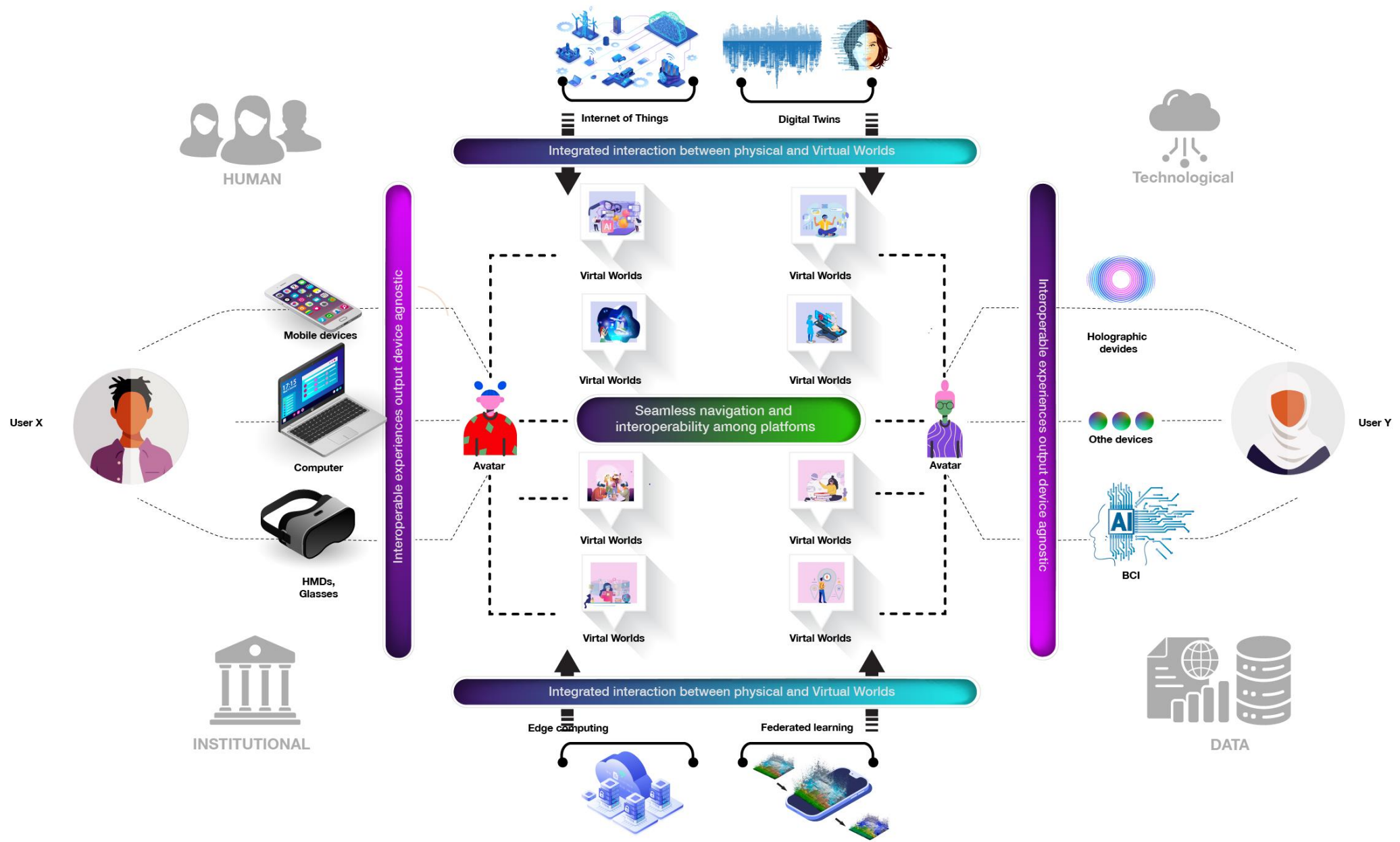
At a broader level, could these changes **impact democratic engagement** and **environmental sustainability?**

Scan to access the XR4HUMAN Code of Conduct for the Human-Centered and Ethical Development of Immersive Technologies

Funded by the European Union

Internet of Things   Digital Twins

Integrated interaction between physical and Virtual Worlds

HUMAN

Technological

Interoperable experiences output device agnostic

Interoperable experiences output device agnostic

User X

Mobile devices

Computer

HMDs, Glasses

Avatar

Virtal Worlds

Virtal Worlds

Virtal Worlds

Seamless navigation and interoperability among platforms

Virtal Worlds

Virtal Worlds

Virtal Worlds

Virtal Worlds

Virtal Worlds

Avatar

Holographic devides

Othe devices

BCI

User Y

Integrated interaction between physical and Virtual Worlds

INSTITUTIONAL

Edge computing   Federated learning

DATA

XR4 HUMAN

Funded by the European Union

XR4 HUMAN

For more info, visit
xr4human.eu

Funded by
the European Union

THANK YOU

FOR YOUR ATTENTION