

Building Trustworthy xWorlds: Privacy, Ethics, and User Agency in AI-Driven Virtual Environments

Yan Hu · Kurt Tutschku

Blekinge Institute of Technology, Karlskrona, Sweden

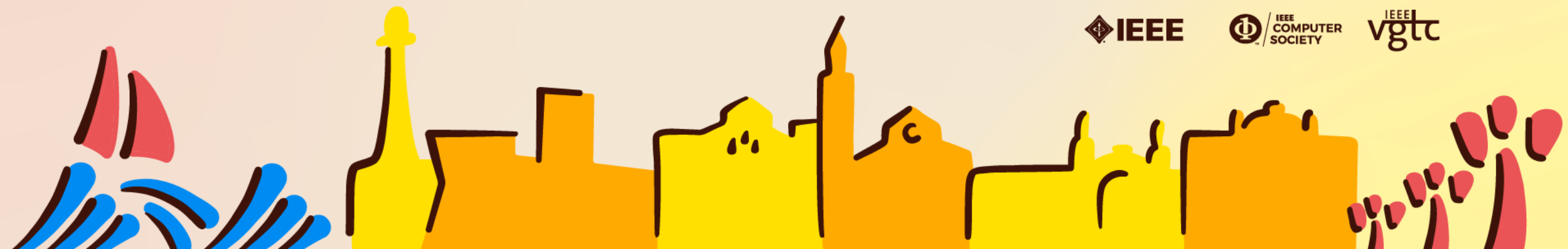
XWORLDS '26 Workshop · ISMAR 2026, Oct 6th, Bari, Italy



ISMAR 2026

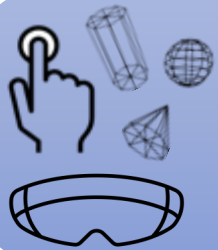
October 5 – 9, Bari, Italy

The 25th International
Symposium on Mixed
and Augmented Reality

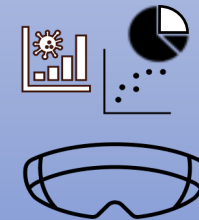




Theme A: Novel Experience
Assessment Methodologies
for Intelligent Realities



Theme B: Novel Environments
and Interaction Techniques
for Intelligent Realities



Theme C: Visual Analytics
for Intelligent Realities



Theme D: Adaptive and
Distributed AI for Intelligent
Realities



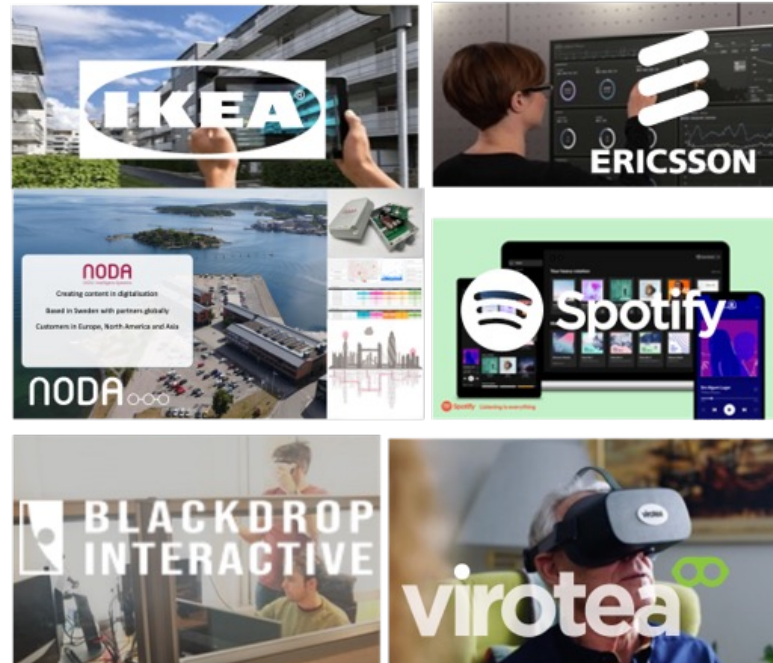
Theme E: Networking for
Intelligent Realities

Human-
Centered
Intelligent
Realities

Budget and Time

- **HINTS profile project funded by the Knowledge foundation Sweden**
- **Six years, 2022-09-01-2028-08-31**
- **Budget 7 million Euro**

Industry Partners



Lab Partner



From isolated worlds to interoperable xWorlds

- ❖ Virtual worlds are becoming persistent, interconnected ecosystems
- ❖ AI, XR and edge computing: agents, humans and physical systems coexist
- ❖ SRIA vision — interoperable, human-centred, aligned with societal values
- ❖ Technical progress is substantial; trust has had far less attention
- ❖ Our position: trust is a design principle, not a by-product



What do we mean by “trust”?

Bursell’s definition

- ❖ *“The assurance that one entity holds that another will perform particular actions according to a specific expectation.”*
- ❖ Not a feeling — a maintained relationship

Three corollaries

- ❖ Contextual — trusted for one action, not for all
- ❖ Temporal — maintained, not granted once
- ❖ Directional — A trusting B says nothing about B trusting A

Interoperability complicates trust

- ❖ Closed platforms keep relationships and data internal
- ❖ Interoperable worlds move data between operators
- ❖ Every transfer introduces a new trust relationship
- ❖ A chain is only as strong as its weakest link
- ❖ So trust cannot be granted once, by one platform

Three pillars, derived from trust itself

1 • Information flow → PRIVACY • data layer

- ⌘ Too much information, and the relationship cannot be monitored

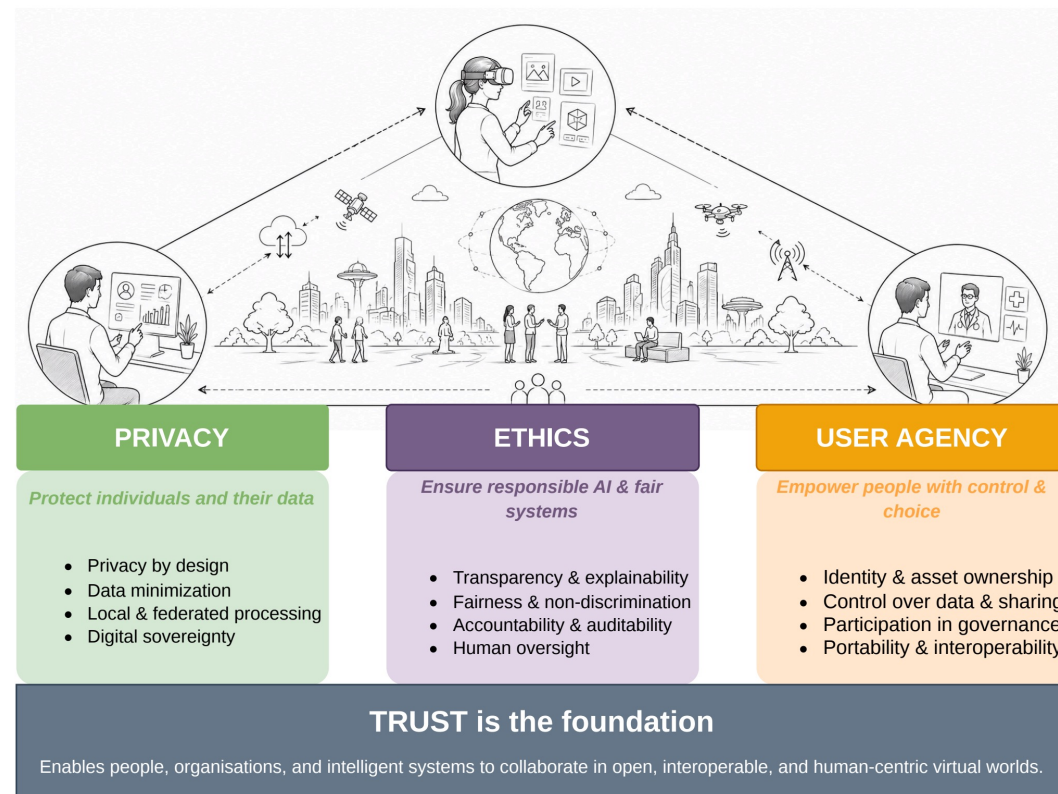
2 • The trustee's actions → ETHICS • system layer

- ⌘ Unobservable behaviour, and expectations become speculative

3 • The trustor's autonomy → USER AGENCY • relationship layer

- ⌘ No way to exit, and trust becomes dependence

A framework for trustworthy xWorlds



Three pillars, three layers

- ❖ Privacy → the data layer
- ❖ Ethics → the system layer
- ❖ User agency → the relationship layer

Trust is a “chemical reaction”

- ❖ It emerges only when all three work together
- ❖ Realised only in open standards
- ❖ **Auditability runs through all three**

“Trust, but verify” – an architectural property

- ❖ The pillars are interdependent, not a checklist
- ❖ Trust that cannot be verified over time is not assurance
- ❖ “Trust, but verify” must be architectural, not aspirational
- ❖ Verification requires open standards
- ❖ Guidance for standards bodies and platform developers

Research challenges

- ❖ Privacy-preserving XR is technologically immature
- ❖ No consent model for inferred and synthetic data
- ❖ Explainability without compromising presence
- ❖ Governance of autonomous virtual agents
- ❖ **Interconnected — progress is needed across all dimensions**

Conclusion

- ❖ From capability-first to trustworthiness by design
- ❖ Trust is an architectural property — not user acceptance
- ❖ Privacy · Ethics · User agency — three interdependent pillars
- ❖ A foundation for researchers, standards bodies, developers, policymakers



Thank you

Building Trustworthy xWorlds: Privacy, Ethics, and User Agency in AI-Driven Virtual Environments



Yan Hu

yan.hu@bth.se · kurt.tutschku@bth.se



Kurt Tutschku