MUNI
FI

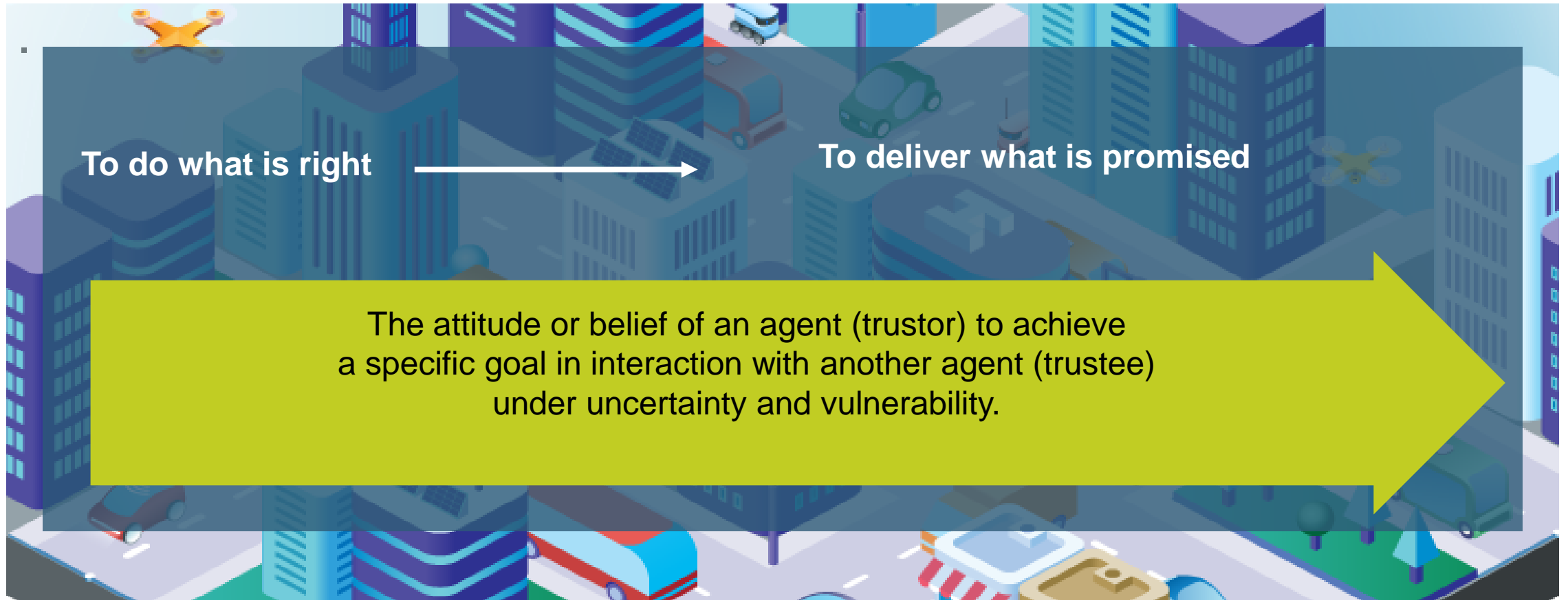# Model-based Approach for Building Trust in Autonomous Systems

Danish Iqbal

Motivation (Autonomous Ecosystem)

— **Multiple, Heterogenous Machines:**

— Varied agent attributes: size, power, mobility, sensing, computing.

— Task allocation based on individual agent capabilities and system goals.

— **Connected (wirelessly):**

— Facilitates coordination and data sharing between agents.

— **Self-Organizing:**

— Can adapt independently to environmental changes.

— Self-organizing ability complements higher-level system directives

MUNI
FI

What is Trust?

To do what is right $\longrightarrow$ To deliver what is promised

The attitude or belief of an agent (trustor) to achieve
a specific goal in interaction with another agent (trustee)
under uncertainty and vulnerability.

D. Iqbal -- Model-based Approach for Building Trust in Autonomous Systems

MUNI
FI

Literature Timeline

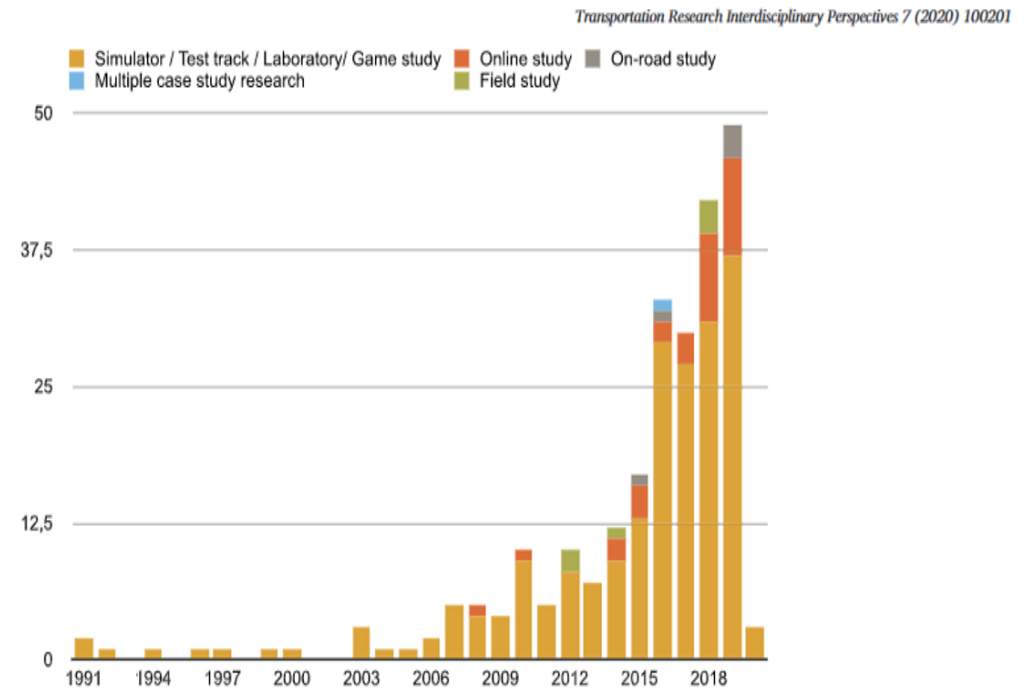## Evolution of Autonomous
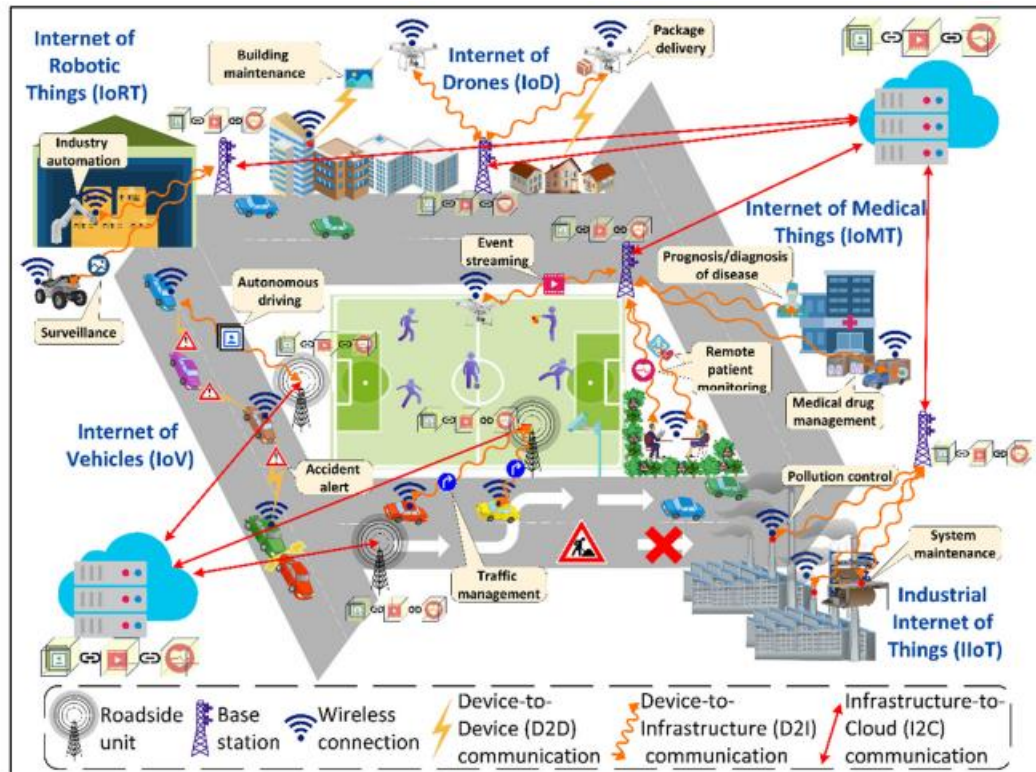
Literature search results categorized into automation, automotive and robotics, A view of smart city digital ecosystem

Why do we need Trust?

Study context chosen for trust research and how they distributed yearly.

[1] Raats, et al. "Trusting autonomous vehicles: An interdisciplinary approach." Transportation Research Interdisciplinary Perspectives 7 (2020): 100201.
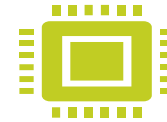
MUNI
FI

Problem Definition

Research Challenges in Autonomous Ecosystem

- — Dynamic and Unpredictable Environments

- — Emergent Behavior (Acquired Knowledge)

- — Insufficient Verification and Validation Methods

- — Lack of Trust Between Systems, Operators, and Networks

- — Dynamic Safety Analysis

D. Iqbal -- Model-based Approach for Building Trust in Autonomous Systems

MUNI
FI

Problem Definition

The "what"

– Goal: Model-based trust assessment in autonomous vehicular ecosystems

– Research questions:

– 1) How can autonomous agents use model-based approaches to dynamically evaluate the trustworthiness of other agents within their ecosystems in real time?

   – **Objective:** Develop a model-based system enabling autonomous agents to dynamically and real-time assess the trustworthiness of collaborating agents in autonomous ecosystem.

– 2) How can a model-based approach identify and react to changes in the behavior of autonomous agents during runtime, especially regarding potential malicious intent?

   – **Objective:** For autonomous systems, in particular, it becomes challenging to dynamically evaluate the predicted behavior of the autonomous agent.

MUNI
FI

Aims of the Research

Chosen Goals

– [O1] **Modelling approach** supporting trust in autonomous vehicular ecosystems

– [O2] **Dynamic safety properties** to be reflected in the model to mitigate collision risks

– [O3] **Reflection of dynamicity and unpredictability of the environment** in which the agents operate

– [O4] **Runtime trust-assessment approach** based on the model (covered by O1 and O2) and the context of the ecosystem (reflecting O3).

MUNI
FI

Proposed Solution

The "how"

– Explore the modeling and assessment technique that supports trust

    – Model-based trust-assessment approach

– **Primary research objectives:**

    1. Modelling approach supporting trust:
       – Modeling Trust Assurance in Ecosystems

    2. **Dynamic safety properties**
       – Dynamic safety analysis of collaborating autonomous agents
       – Dynamicity is difficult to achieve trust

– **Proposed solutions**

The concept of Digital Twin (DT)
- Role of DTs in supporting runtime evaluation of autonomous agents
- Design a model that allows autonomous systems to share their declared behavior in the form of a Digital Twin, which is used for trust evaluation in dynamic environments
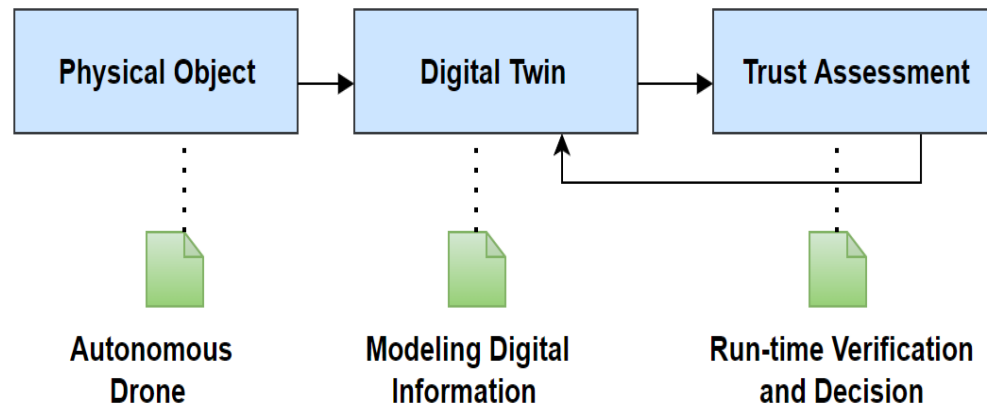
Safety mechanism integration
- Trust assessment through DT provided by collaborators
- Checking of actual and declared behaviour

MUNI
FI

Proposed Solution

The "how"

– Explore the modeling and assessment technique that supports trust

    – Model-based trust-assessment approach

– **Proposed solution**

Elements of the Proposed Solution

– **Research contributions**

– Concept of Digital Twin (DT) and trust assessment (run-time verification)
– Introduce a modeling and assessment technique that supports trust-building in autonomous vehicular ecosystem.
– Suggest runtime verification method that facilitates trust assessment through information exchange, utilizing a Digital Twin and conducting real-time assessments of this Digital Twin.

MUNI
FI

Proposed Solution

The "how"

— Explore the modeling and assessment technique that supports trust

— Safety mechanism integration

— **Proposed solution**
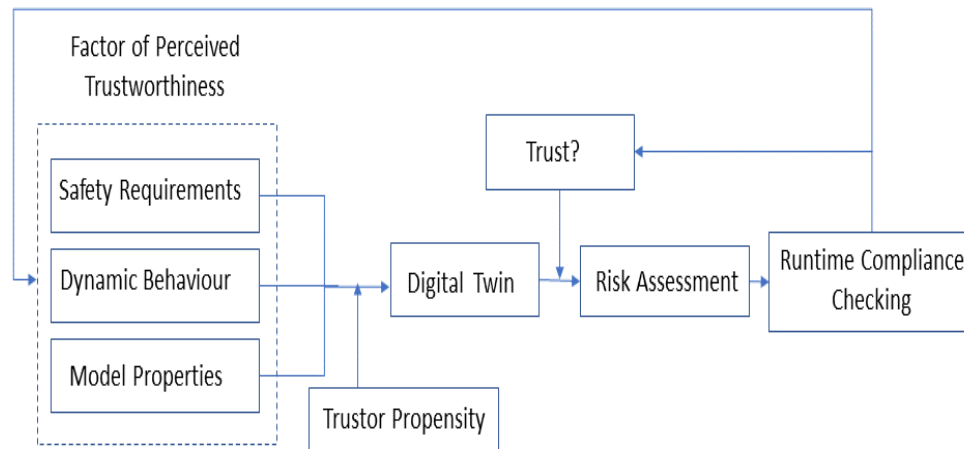
Trust Modeling through Digital Twin

— **Research contributions**

— Safety-based trust-building method

— This is done to reflect two major social metrics relevant to trust building, which are:

— 1) Honesty (consistency between the declared and actual behaviour)

— 2) Openness (transparency about the intended behaviour).

MUNI
FI

Achieved results

## Model-based Approach for Building Trust:

- The first version of envisioned approach to model-based trust assurance in autonomous vehicle ecosystems, contributing to O4, along with the first Digital Twin model, contributing to O1, has been published in [1].

- [1] Model-based Approach for Building Trust in Autonomous Drones through Digital Twins
    IQBAL, Danish; BUHNOVA, Barbora. (SMC). 2022. | (CORE Rank: A)

Drone 1 Decision to Trust Drone 2

MUNI
FI
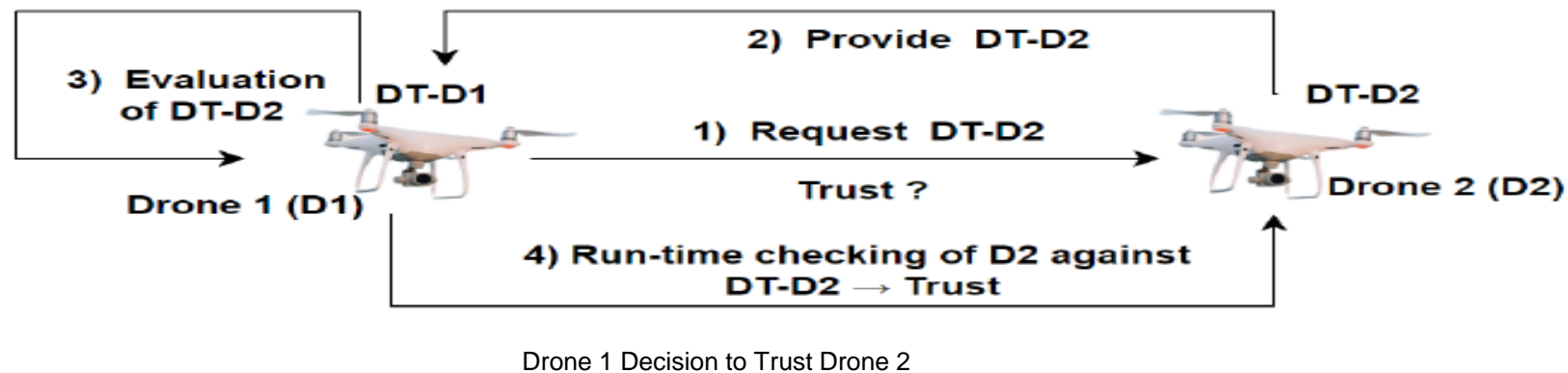
Achieved results

Trust Building through Dynamic Safety Evaluation:

‒ Introduced a dynamic trust assessment approach for collaborative autonomous systems.

‒ Approach incorporates dynamic safety assurance to contribute to Objective O2.

‒ Utilizes Digital Twins for real-time data exchange.
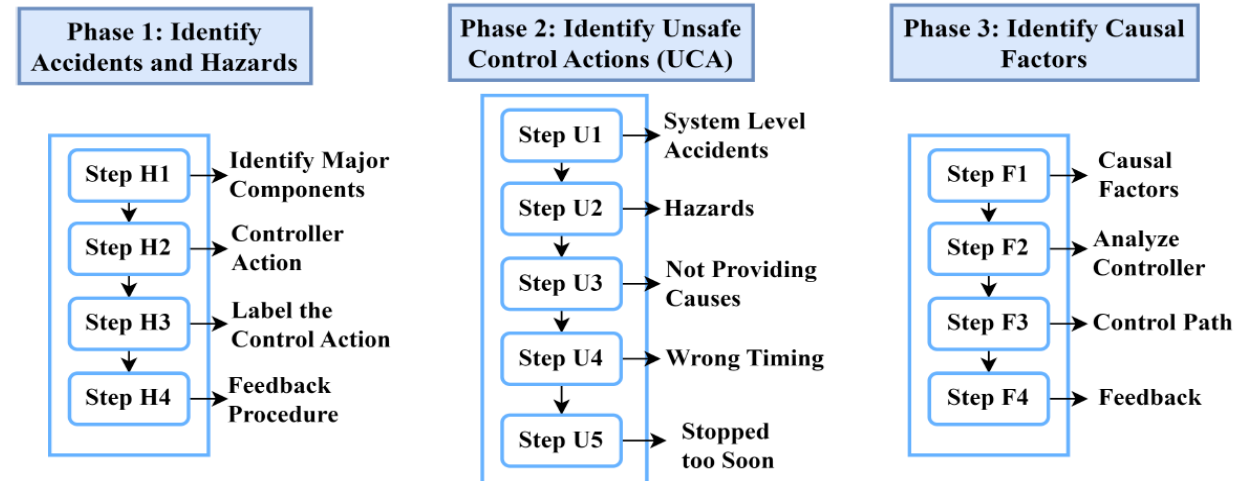
Framework for Identification of Hazard, Unsafe Control Action and Casual Factors.

‒ [2] Digital Twins for Trust Building in Autonomous Drones through Dynamic Safety Evaluation
IQBAL, Danish;BUHNOVA, Barbora; CIOROAICA, Emilia. 2023 | Enase (CORE Rank: B)

MUNI
FI

Conclusion

– Examined the role of Digital Twins in the trust-building process.

– Integration of safety mechanism and run-time compliance checking.

– Direct and indirect trust on the practical use case.

– Finally, we presented the trust/not trust example for collision avoidance.

– In the future, we plan to improve the current design of the Digital Twin by including more detailed behavior, presenting an enhanced model.

– we plan to employ the trust assessment to validate the model with more attributes and scenarios.

D. Iqbal -- Model-based Approach for Building Trust in Autonomous Systems

MUNI
FI