SCIENCE
PASSION
TECHNOLOGY

Trusted AI at TU Graz

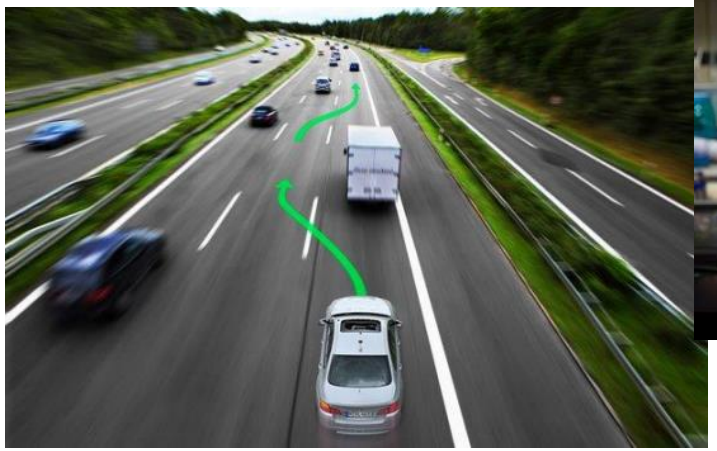
Bettina Könighofer
bettina.koenighofer@tugraz.at

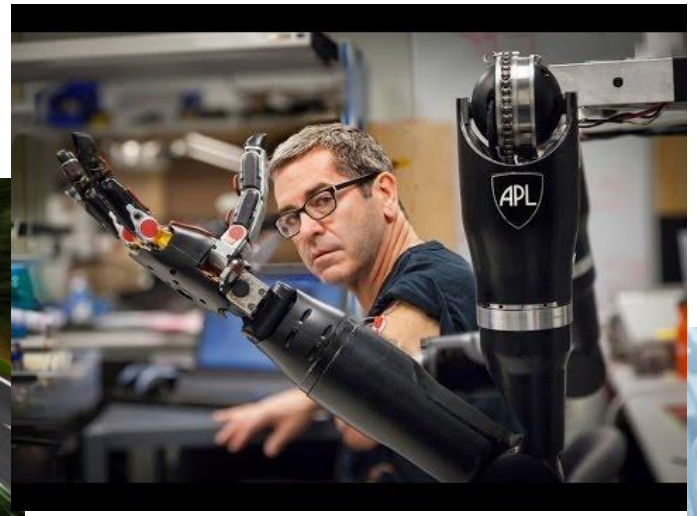
July 23, 2024

Design Turstworthy AI-based Systems

Safety | Explainability | Accountability | Fairness

Robustness | Sustainability | Security | Privacy





https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Design Turstworthy AI-based Systems

Safety | Explainability | Accountability | Fairness

Robustness | Sustainability | Security | Privacy

Bernhard Aichernig

Franz Wotawa

Roman Kern

Bettina Könighofer

Roderick Bloem

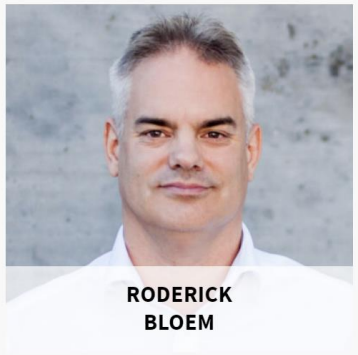
Christian Rechberger

Elisabeth Lex

Martin Tappler

Institute for Security (IAIK)


STEFAN MANGARD


RODERICK BLOEM


MARIA EICHLSEDER


DANIEL GRUSS


BETTINA KÖNIGHOFER

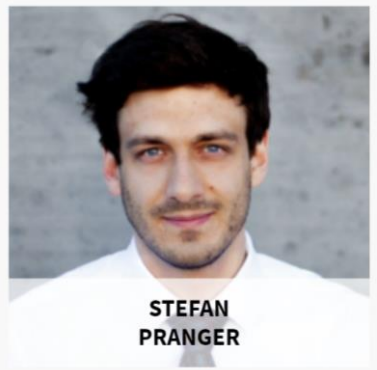

CHRISTIAN RECHBERGER

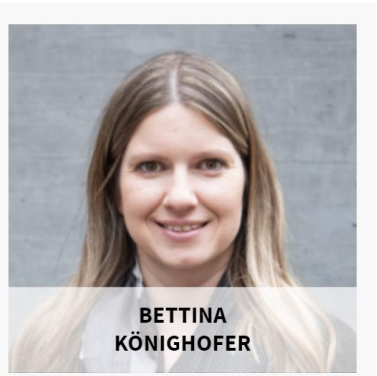

SUJOY SINHA ROY

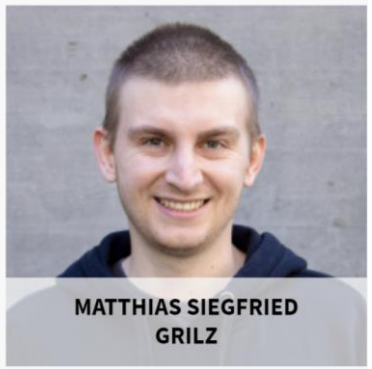

REINHARD POSCH

Trusted AI Group at IAIK

Safety | Explainability | Accountability | Fairness | Robustness

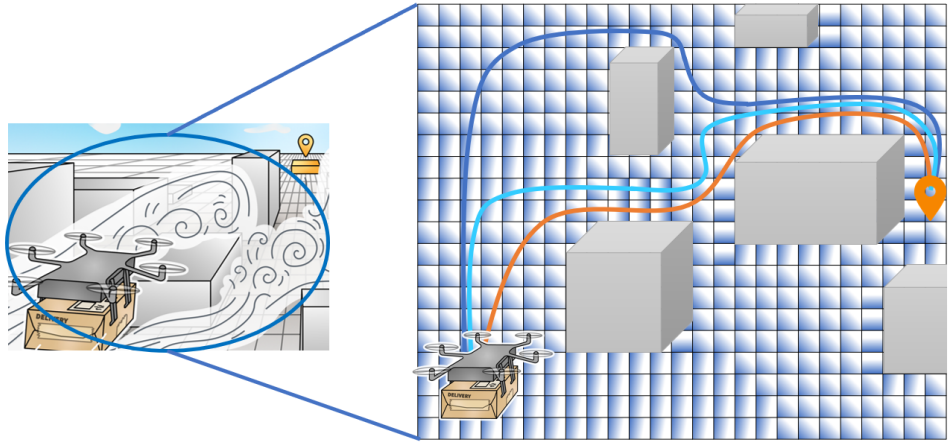


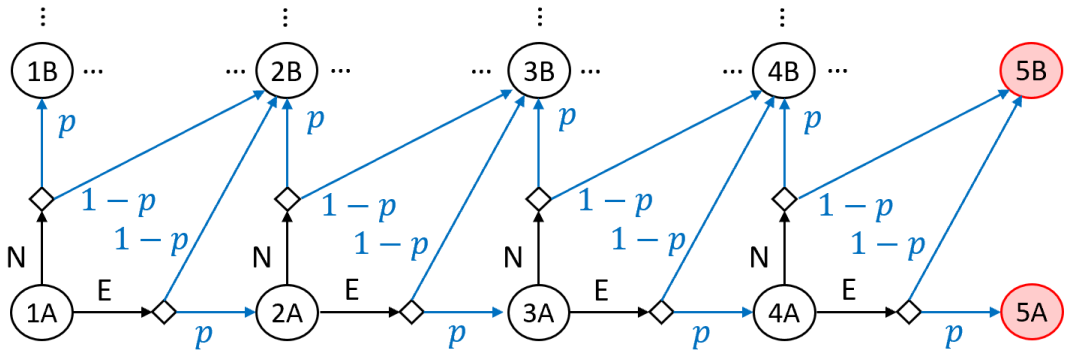




Combine Symbolic AI with Data-Driven AI

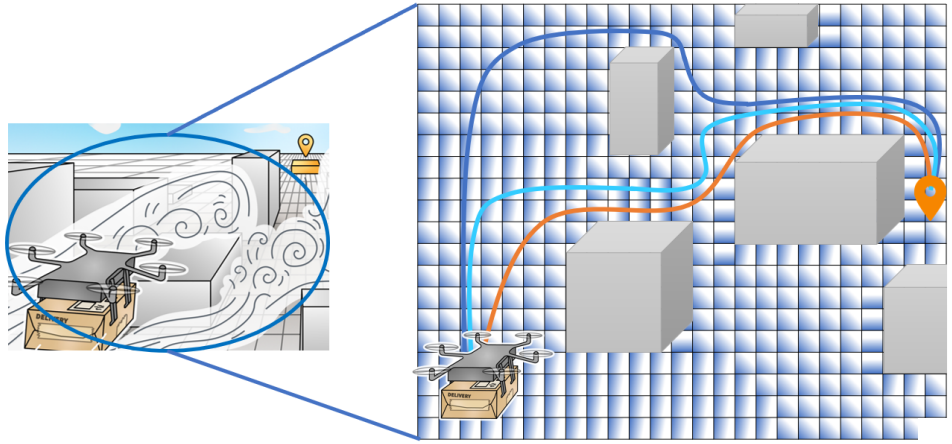
Model system behaviour

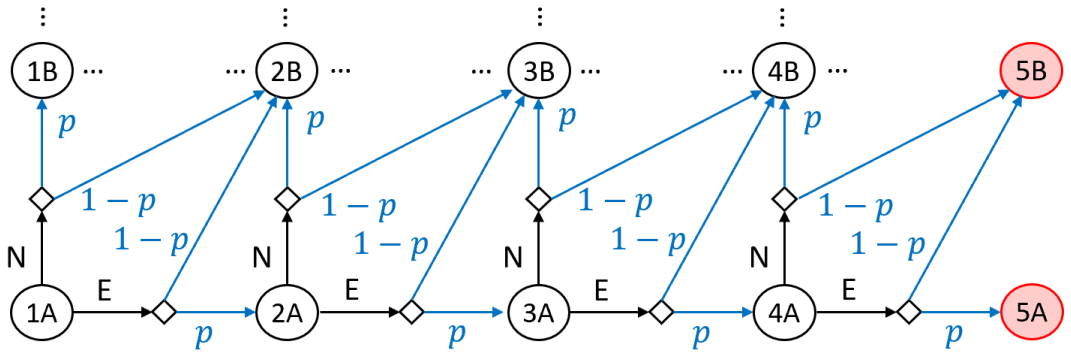




Combine Symbolic AI with Data-Driven AI

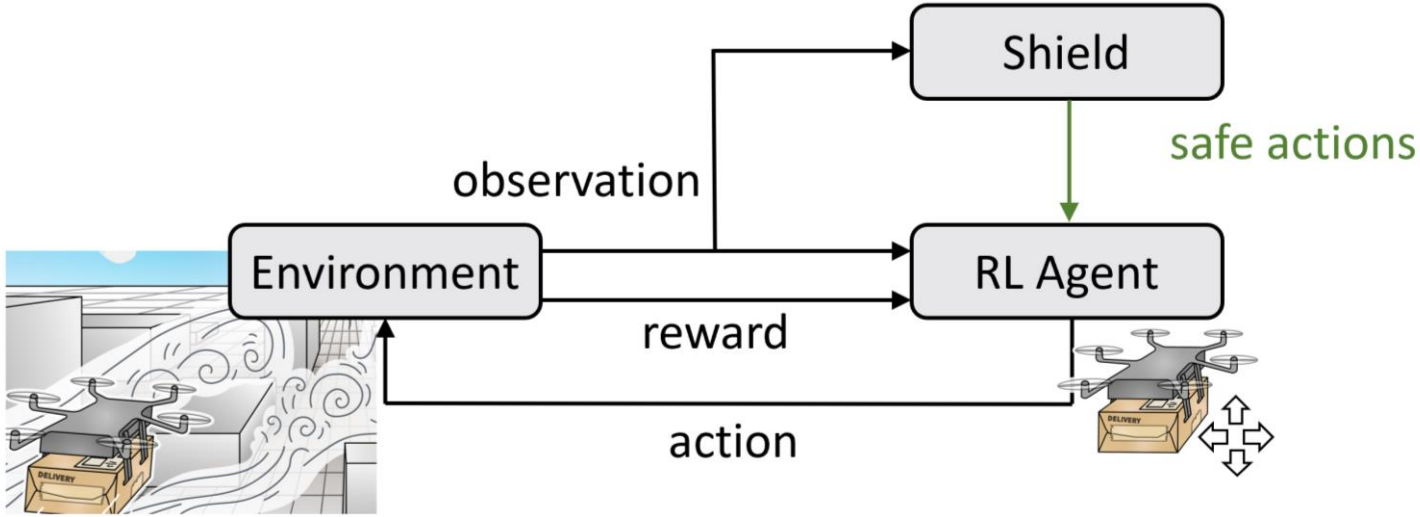
Model system behaviour





Use model

- Testing, Verification
- Enforcing Safety
- Monitoring Fairness
- Explainabiltiy
- Risk Analysis…



Thank you!

Contact me for collaborations

- Joint project proposals
- Fair and correct decision making
- Fundamental research / industrial applications

Bettina Könighofer
bettina.koenighofer@tugraz.at