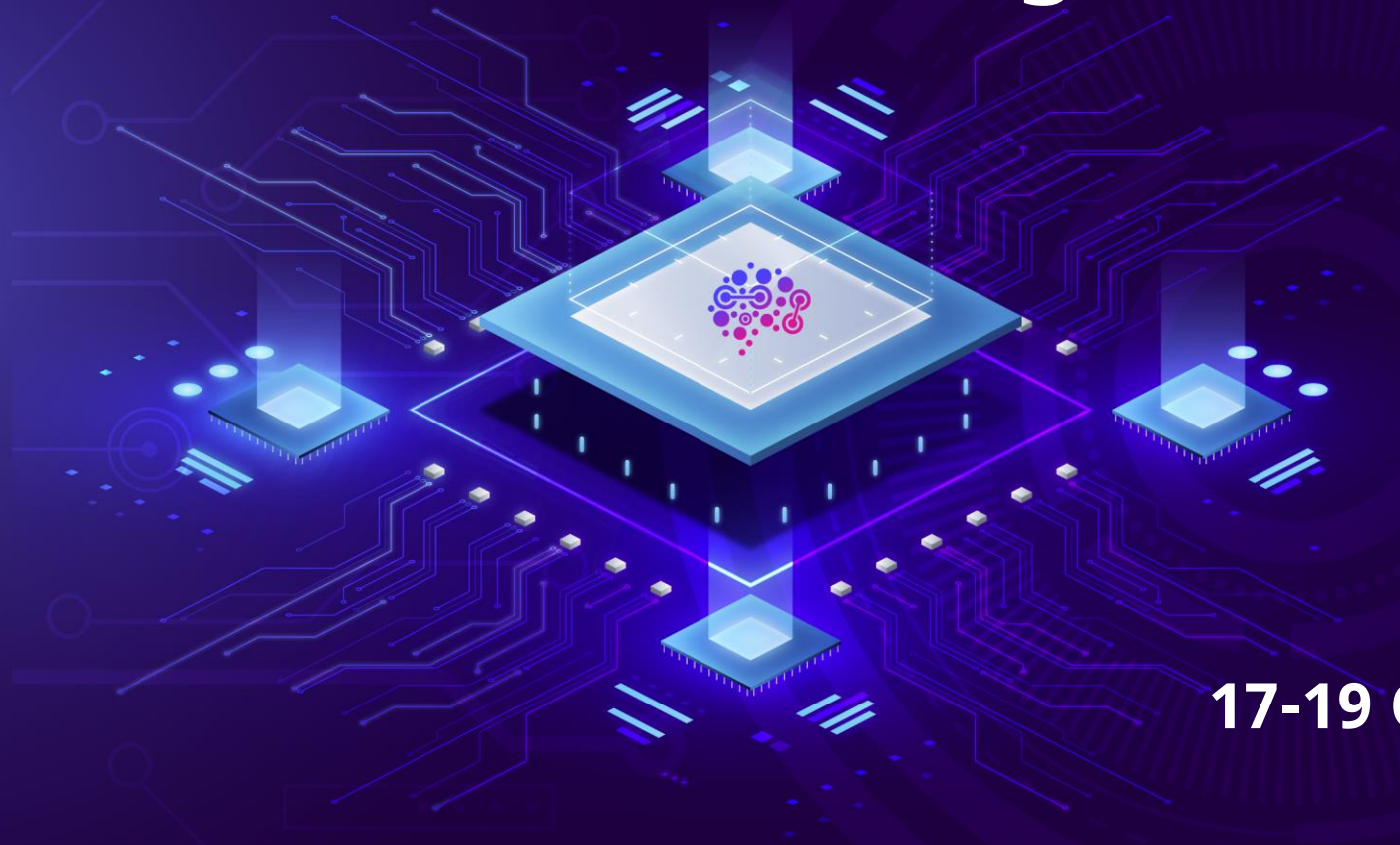
European Conference on EDGE AI Technologies and Applications - EEAI



Advancing emerging edge AI technologies and driving next generation intelligent applications.

17-19 October 2023, Athens, Greece

European Conference on EDGE AI Technologies and Applications - EEAI

Shaping Edge AI that flows with Humans: Designing Trustworthy and Useful AI Applications

Nikolai Ebinger

Peter Moertl

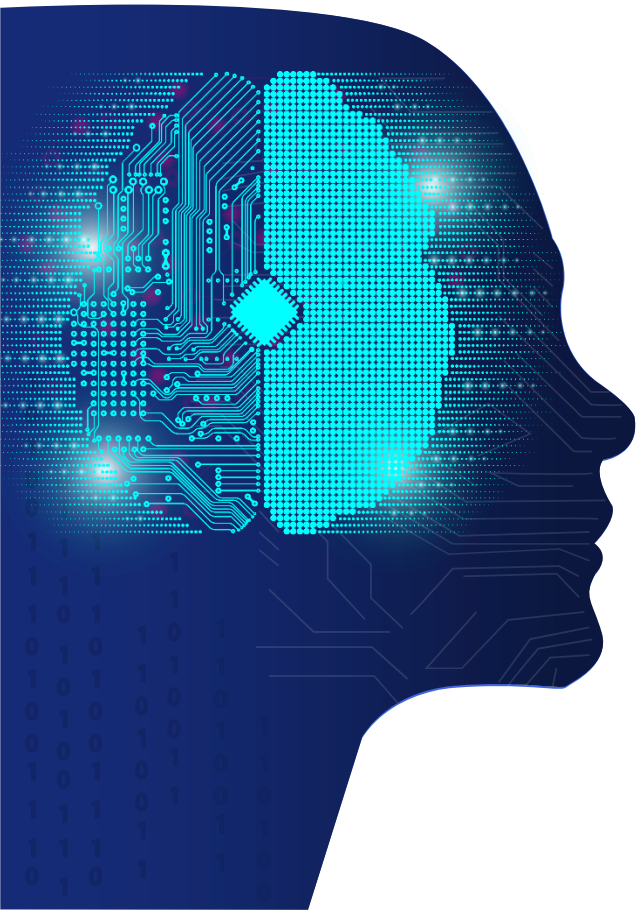
Michael Karner

Virtual Vehicle Research GmbH
Graz, Austria

17-19 October 2023 Athens, Greece

Copyright © 2023

Presentation Outline



- AI developments are blossoming
- Why ethics and trust are necessary for acceptance and operational uptake of AI
- Our framework to design AI for acceptance, ethics and trust

Copyright © 2023

AI developments are blossoming

- Increased productivity
  - Taking over repetitive tasks
  - Allowing humans time for strategic & creative tasks
  - AI as helper
- Improved decision making
  - Turn complex data into digestable insights
  - Isolate trends
  - Identify changes
  - Predictive analysis
- Improved safety
  - Taking over hazardous missions
  - Increase data and information security

https://th.bing.com/th/id/OIP.Vx0HDTgZpz-gc2xTZbCb_AHaEK?pid=ImgDet&rs=1

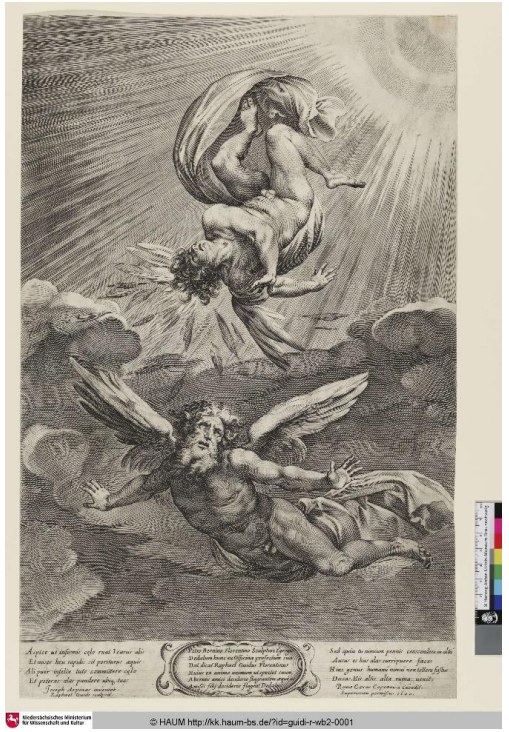
https://figes.com.tr/wp-content/uploads/2021/01/GettyImages-993491068-scaled-1.jpg

https://www.beekeeper.io/blog/3-reasons-you-want-ai-in-the-workplace/
https://www.linkedin.com/pulse/impact-artificial-intelligence-ai-workplace-charlotte-volet/

Copyright © 2023

The role of ethics and trust for acceptance and operational uptake

- The promises of AI are such that businesses have to adopt AI
- However, there are backlashes
  - 2023 Hollywood strikes are also about AI
    - Use of generative AI
  - 2023 demise of a small company Prosecraft
    - „Prosecraft's founder didn't see the controversy coming."
- Wall Street Journal, Aug 2023:
  - "More and more companies say they are concerned about facing public criticism over their use of artificial intelligence, thanks to rising fears over the technology's negative impacts, including job losses."



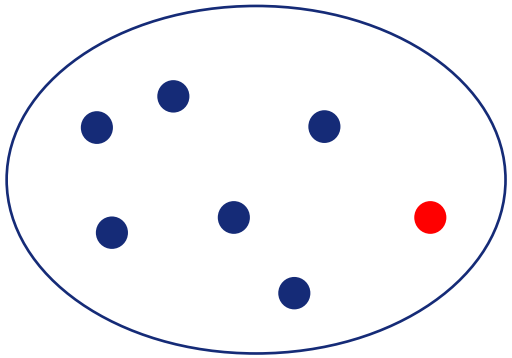
https://nat.museum-digital.de/singleimage?resourcenr=1064875

https://www.wired.com/story/prosecraft-backlash-writers-ai/
https://www.wsj.com/articles/companies-increasingly-fear-backlash-over-their-ai-work-53aff47c https://www.vulture.com/article/wga-strike-2023.html

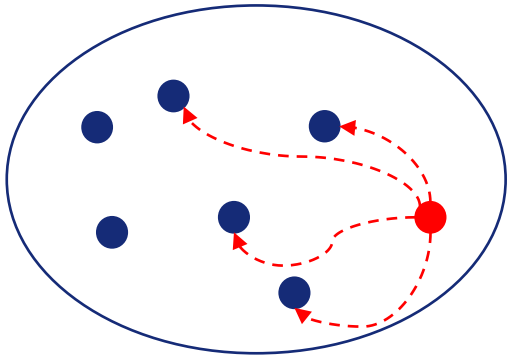
Copyright © 2023

Why it is inherently difficult to design AI for acceptance, ethics and trust

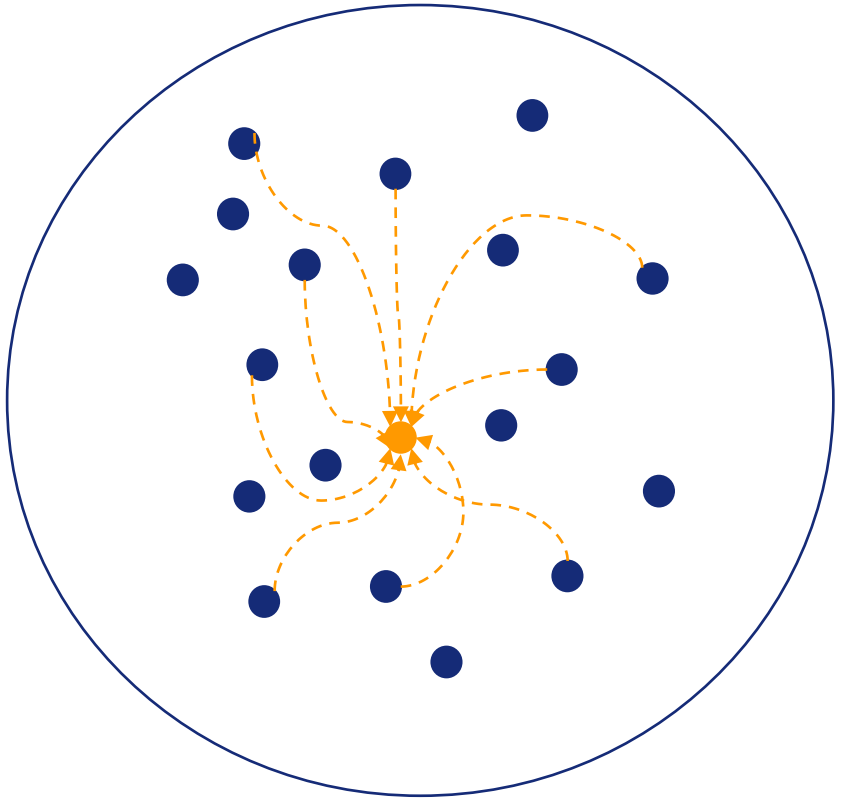
A new AI application starts often with a technical innovation idea: algorithms and data



At this time, the use conditions, users, organizations, tasks, and responsibilities are often unknown to the developers



This can lead to non-acceptance by users, mistrust, or even the above Societal backlashes

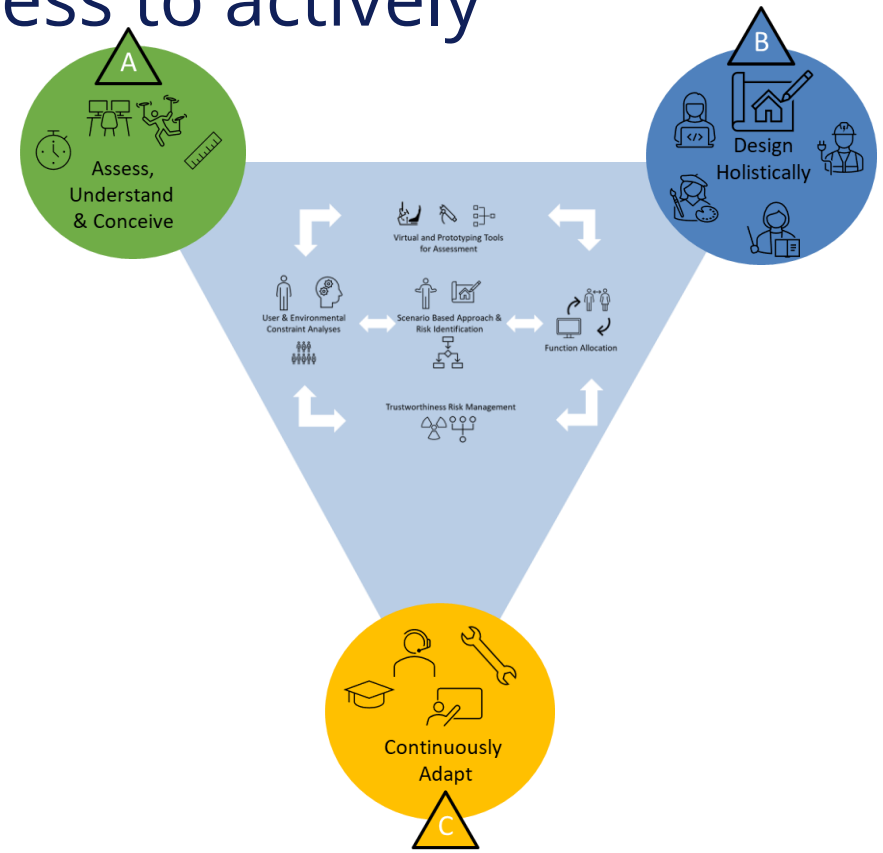


The challenge is to design the AI application with the context (users, organizations, tasks, responsibilitiies in mind)

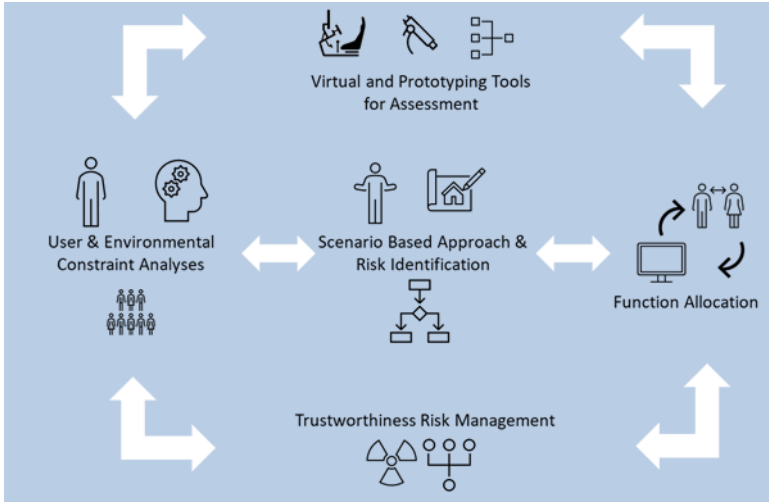
Copyright © 2023

Human Systems Integration Framework

As part of the InSecTT project have we developed a framework and process to actively



Minimum Ecosystem to develop trustworthy systems

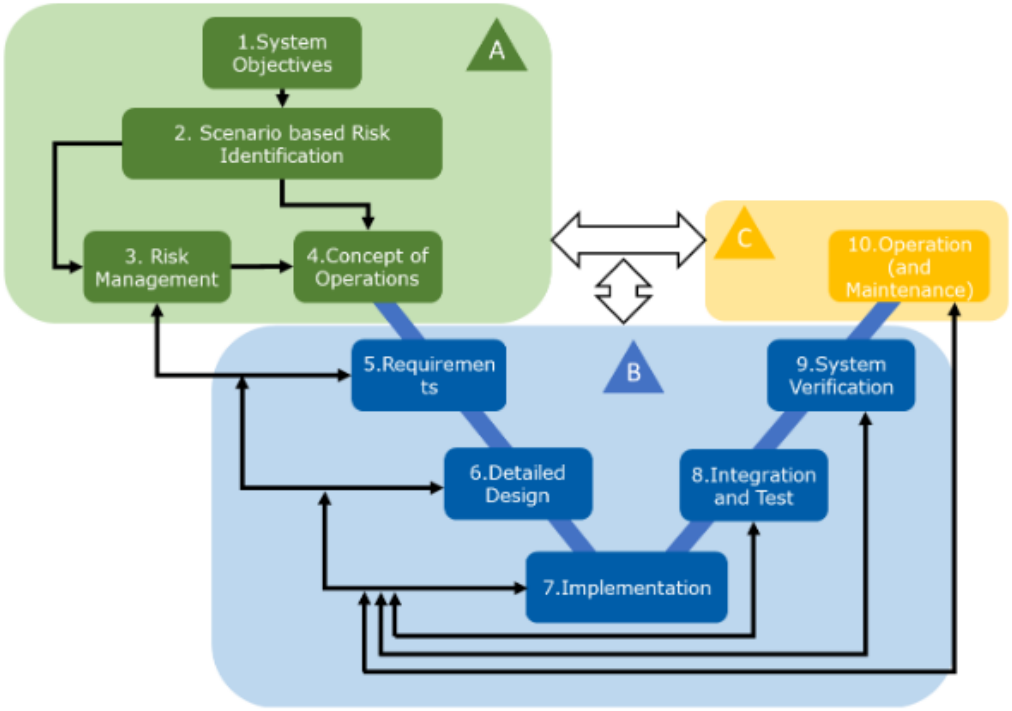


Basic process and tools

Copyright © 2023

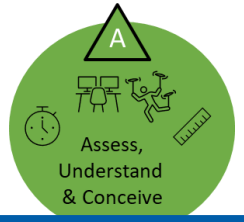
Applications of the HSI Framework

- To make this more concrete:
  - As part of the InSecTT project we applied this framework to diverse use cases
    - Transitioning from automated driving back to manual driving
    - Automated detection of security problems in manufacturing environment
    - *Emergency and rescue operations*
    - *Hospital bed management*
- Our Conclusions



Copyright © 2023

Example: Holistic approach to partial driving automation



Literature Review

Task Analysis

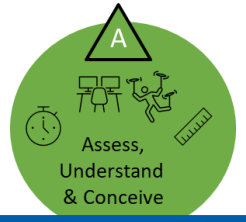
Identified Challenges
- System functionality & System limitations,
- Driver responsibility / division of tasks,
- Transition Procedure,
- System Warnings and Information,

User Study on the state of the Art

Identified Challenges
- System functionality & System limitations
- Mode confusion
- General interaction
- Anticipation system behavior
- Over-/Undertrust

Copyright © 2023

Example: Holistic approach to partial driving automation



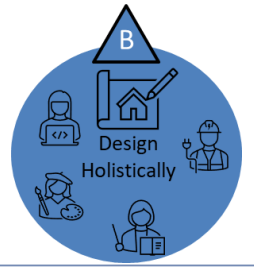
Literature Review

Task Analysis

Identified Challenges
- System functionality & System limitations,
- Driver responsibility / division of tasks,
- Transition Procedure,
- System Warnings and Information,

User Study on the state of the Art

Identified Challenges
- System functionality & System limitations
- Mode confusion
- General interaction
- Anticipation system behavior
- Over-/Undertrust



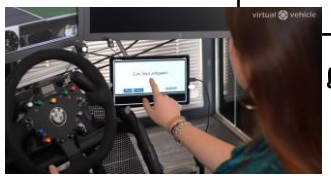
Holistic Design Expands Driving Automation System

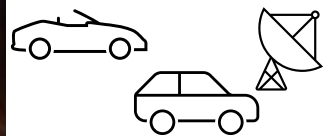
Environment
Road Infrastructure
Sensors and Messages

Vehicle Cabin
Driver State Monitoring
and Adaptive Interactions

Driver
Help Driver to meet
new responsibilities

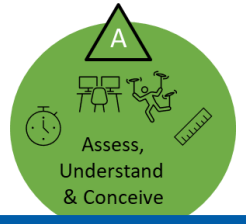
Automation Predictability





Copyright © 2023

Example: Holistic approach to partial driving automation



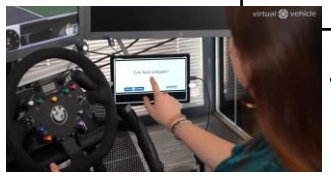
Literature Review

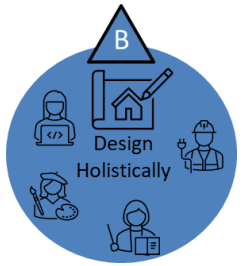
Task Analysis

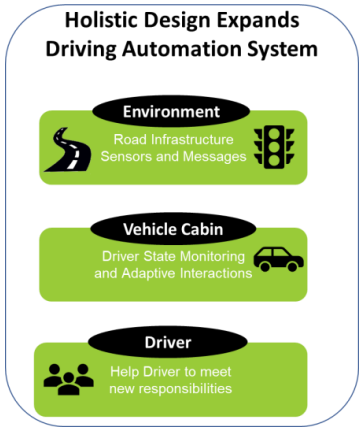
Identified Challenges
- System functionality & System limitations,
- Driver responsibility / division of tasks,
- Transition Procedure,
- System Warnings and Information,

User Study on the state of the Art

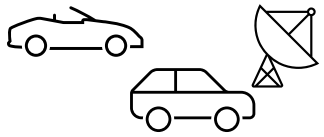
Identified Challenges
- System functionality & System limitations
- Mode confusion
- General interaction
- Anticipation system behavior
- Over-/Undertrust

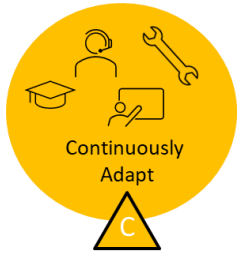






Automation Predictability



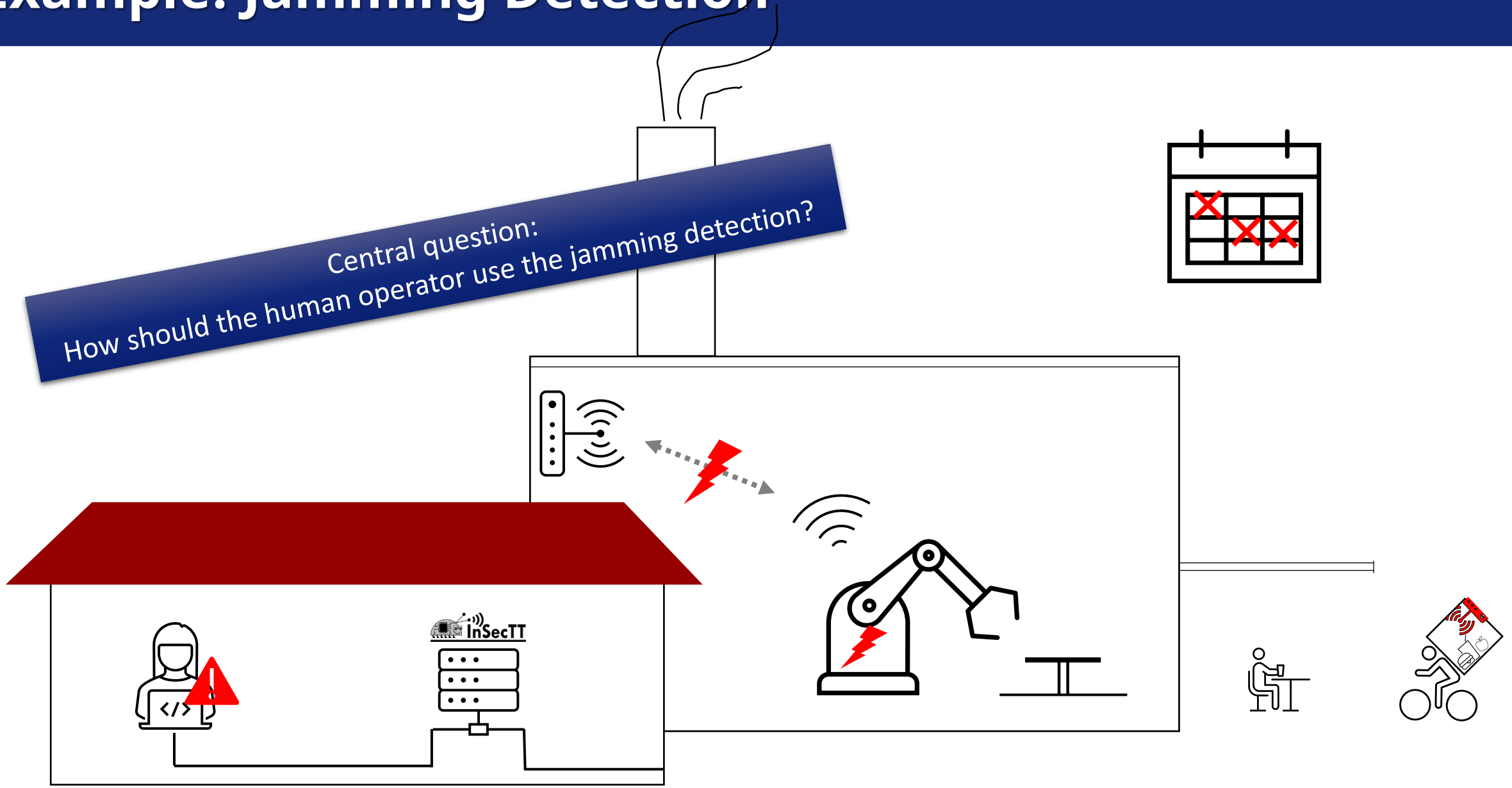


Next Step:
Eco System

Copyright © 2023

Example: Jamming Detection

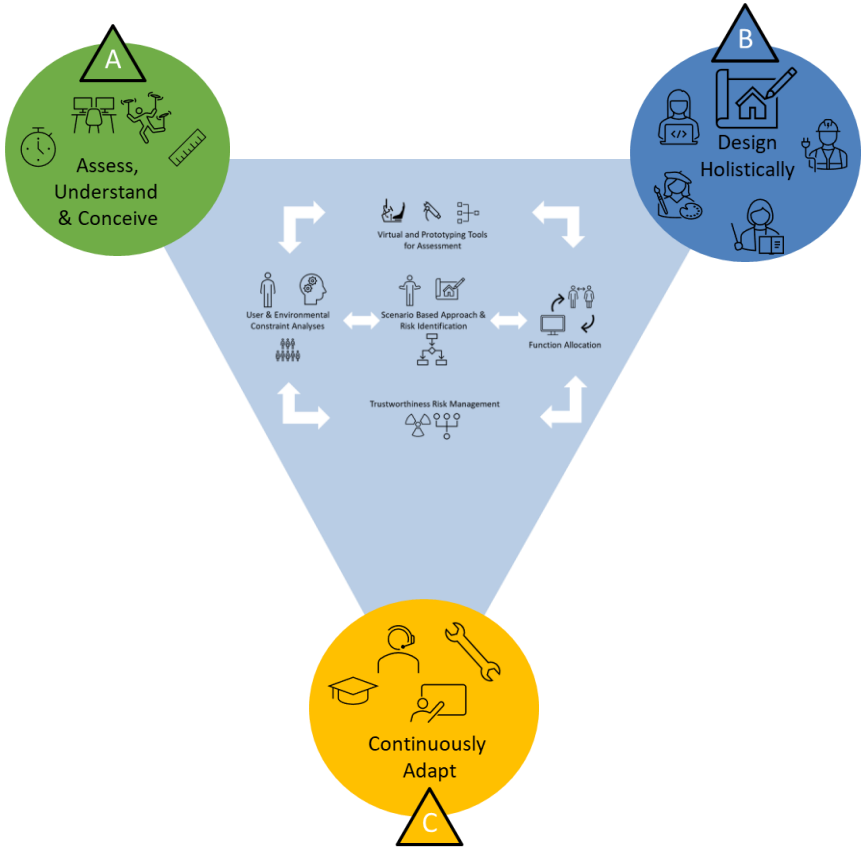
Central question:
How should the human operator use the jamming detection?



Copyright © 2023

Conclusion

- Existing technical processes are surprisingly resistant against designing trust and acceptance
  - Silos of technical excellence
- A holistic approach and eco-system seem necessary to break them open and shape technology for humans
- This requires sufficient contextualization from the beginning of the AI development
- Tight interactions between inter-disciplinary teams
- New tools and processes for development and risk management
- Extended product life-cycles → increased complexity!



Copyright © 2023

Conclusion

- Existing technical processes are surprisingly resistant against designing trust and acceptance
  - Silos of technical excellence
- A holistic approach and eco-system seem necessary to break them open and shape technology for humans
- This requires sufficient contextualization from the beginning of the AI development
- Tight interactions between inter-disciplinary teams
- New tools and processes for development and risk management
- Extended product life-cycles → increased complexity!

At this conference:
Poster on „Human Oversight“

Whitepaper:

Copyright © 2023

Copyright © 2023

EDGE AI

virtual vehicle

Thank You

For your attention

@ nikolai.ebinger@v2c2.at

Federal Ministry Republic of Austria Climate Action, Environment, Energy, Mobility, Innovation and Technology

Federal Ministry Republic of Austria Labour and Economy

FFG Promoting Innovation.

Das Land Steiermark Wirtschaft, Tourismus, Wissenschaft und Forschung

SFG

Virtual Vehicle Research GmbH has received funding within COMET Competence Centers for Excellent Technologies from BMK, BMAW, the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management.

The project HOFAI (FFG) was funded by the program "Industrial Dissertation" sponsored by Nationalstiftung für Forschung, Technologie und Entwicklung and Österreich-Fonds, supported by the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK).

InSecTT project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876038. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Austria, Sweden, Spain, Italy, France, Portugal, Ireland, Finland, Slovenia, Poland, Netherlands, Turkey

Disclaimer excluding JU responsibility: The document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.