German Research Center for Artificial Intelligence (DFKI)



Trusted AI in Deutschland

Philipp Slusallek

German Research Center for Artificial Intelligence (DFKI) Saarland University, Computer Graphics Lab (UdS-CGL)



Saarland Informatics Campus slusallek@dfki.de

Al systems literally build themselves...

"self-created"





- No "programmer's intention"
- No guarantees of *what* has been learned
- No guarantees of performance
- No explanations for humans



... and are connected in systems of exponentially growing complexity









... and are connected in systems of exponentially growing complexity









We are reaching the limits of control over Al systems





Limited understanding of

- Functionality
- Integrity
- Reliability
- Safety
- Generalization





Traditional engineering fails for AI (1/2)



Two approaches for quality assurance







Traditional engineering fails for AI (2/2)



Two approaches for quality assurance







... leading to dwindling public acceptance and loss of markets

Waymo in May 2018: "Up to 62,000 autonomous taxis by end of 2018"







2021: Still no commercial autonomous driving Uber and others even stopped development entirely





"Pushing the limits" of controllability – trust as value proposition!









Closing the gap requires a concerted effort of society as a whole





Imperatives

- Acknowledge importance
 and urgency
- Seize the opportunity for differentiation
- Act decisively, aim high
- Move from incremental to disruptive innovation





Closing the gap requires a concerted effort of society as a whole





Imperatives

- Acknowledge importance
 and urgency
- Seize the opportunity for differentiation
- Act decisively, aim high
- Move from incremental to disruptive innovation





Germany and Europe are best positioned to achieve this





Assets

- Long-standing social/ethical heritage
- Liberal, responsible economic policy
- Quality leadership and attribution
- High innovation rate
- Scientific excellence and thoroughness
- Strong AI research community





Possible Approach to Trusted Al



SIC Saarland Informatics Campus







CLAIRE Germany

- Grow as a community in Germany and connect with the rest of Europe
 - Closer collaboration within research groups (\rightarrow Research network)
 - Closer collaboration with industry (\rightarrow Innovation network)
 - Identify specific interests for Germany ecosystem
- Drive selected AI topics relevant for Germany
 - Trusted AI?
 - Large-Scale European AI Models (LEAM, Open GPT-X)?
 - Common AI platform for Gaia-X?
 - Others?
- CLAIRE Germany is a great community to discuss this





Thank you very much for your attention !







