



Wolfgang Zankl, University of Vienna

EU Ethics Guidelines for Trustworthy Artificial Intelligence

Tianjin, May 16th, 2019

A. Need for Specific Regulation?

- No, not now - too early
- Bad example: Electronic Persons - EU Parliament Resolution with recommendations to the Commission on Civil Law Rules on Robotics, February 16th 2017.
- Regulation at an early stage might obstruct innovation
- AI: no concrete plans to write up legislation
- Instead: update existing regulation in areas such as safety and data protection by taking into account ethical guidelines

B. Need for Specific Guidelines?

- Yes, together with stakeholders
- Good example: EU AI strategy - COM (2018) 237, COM (2018) 795
- Leading to Communication from the EU Commission „Building Trust in Human-Centric Artificial Intelligence“, April 8th, 2019 - COM (2019) 168 final
- Based upon „Ethics Guidelines for Trustworthy AI“ by High-Level Expert Group on Artificial Intelligence set up by the European Commission:

C. Ethics Guidelines for Trustworthy AI

- Aim: promotion of trustworthy AI
- **Three components** of trustworthy AI:
 1. Lawful
 2. Ethical
 3. Robust
- Guidelines
 - focus on 2 and 3
 - are non-binding and as such do not create any new legal obligations
 - apply to all AI systems, but specific context should be taken into account for their concrete and proportionate implementation: AI suggesting unsuitable books vs. AI misdiagnosing cancer

- ***Seven key requirements***, that AI applications should respect to be considered trustworthy:

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Societal and environmental well-being
7. Accountability

D. Next steps

2019: Piloting phase involving stakeholders

2020: Building on feedback guidelines will be reviewed and updated



Thank you for your attention!

wolfgang.zankl@univie.ac.at