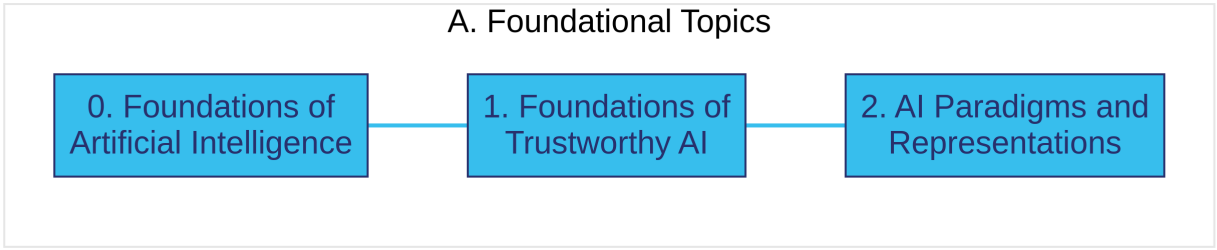
Contents

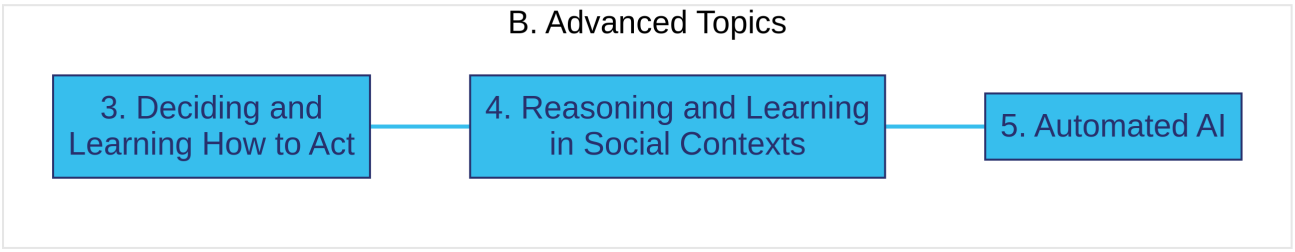
Foundations of Trustworthy AI

> **This website is currently a demo**
>
> The content of this website is a demonstrator of the functionalities that could be used in a MOOC.

This website serves as an example of the possible online content for a Massive Open Online Course aligned with the TAILOR PhD curriculum in Trustworthy AI.

The PhD curriculum in Trustworthy AI equips students with essential knowledge and skills for developing trustworthy Artificial Intelligence systems. With a general foundation on Artificial Intelligence, students learn how an integrated approach to learning, optimisation and reasoning can achieve trustworthiness. The first three topics are introductory, with (0) a general overview of Artificial Intelligence that ensures a common understanding across the curriculum; (1) a special focus on the trustworthy aspects of AI; and (2) how to leverage complementary AI paradigms and representations. The last three advanced topics include (3) AI agents deciding and learning how to act; (4) AI agents acting and learning in society; and (5) ensuring that AI tools and systems are performant, robust and trustworthy.





The content of this MOOC focuses on the first topic proposed in the TAILOR PhD curriculum **Foundations of Trustworthy AI**. The idea is to adapt the content of The TAILOR Handbook of Trustworthy AI and create a course with the following subtopics:

1. Explainable AI Systems
2. Safety and Robustness
3. Fairness, Equity, and Justice by Design
4. Accountability and Reproducibility
5. Respect for Privacy
6. Sustainability