Framework for Trustworthy AI

Trustworthy AI

Lawful AI	Ethical AI	Robust AI

(not dealt with in this document)

Foundations of Trustworthy AI

Adhere to ethical principles based on fundamental rights

→ **4 Ethical Principles**

Acknowledge and address tensions between them

→
- Respect for human autonomy
- Prevention of harm
- Fairness
- Explicability

Realisation of Trustworthy AI

Implement the key requirements

→ **7 Key Requirements**

Evaluate and address these continuously throughout the AI system's life cycle via

↙ **Technical Methods** ↘ **Non-Technical Methods**

→
- Human agency and oversight
- Technical robustness and safety
- Privacy and data governance
- Transparency
- Diversity, non-discrimination and fairness
- Societal and environmental wellbeing
- Accountability

Assessment of Trustworthy AI

Operationalise the key requirements

→ **Trustworthy AI Assessment List**

Tailor this to the specific AI application