Erasmus Centre for Data Analytics
Expert Practice *Trustworthy & Accountable AI*

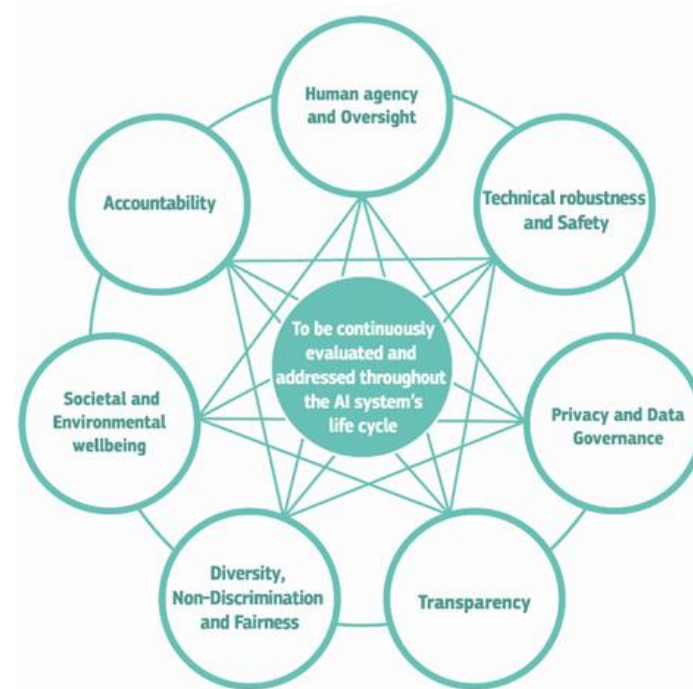# Trustworthy & Accountable AI: Introduction

Dr. Iuliana Sandu, Director Expert Practice Trustworthy & Accountable AI

"Metrics play a central role in Artificial Intelligence (AI) applications. The design of metrics, the impact of metrics, and the use of metrics with the purpose of upholding accountability and ultimately increase trust, has been a long endeavor of researchers in accounting and management information systems. In the Trustworthy and Accountable AI Expert Practice, part of the Erasmus Center for Data Analytics (ECDA), AI is looked at through the lenses of experts on accountability and metrics."

Trustworthy & Accountable AI: Context

INDEPENDENT
HIGH-LEVEL EXPERT GROUP ON
ARTIFICIAL INTELLIGENCE
SET UP BY THE EUROPEAN COMMISSION

ETHICS GUIDELINES
FOR TRUSTWORTHY AI

Accountability

Human agency and Oversight

Technical robustness and Safety

To be continuously evaluated and addressed throughout the AI system's life cycle

Privacy and Data Governance

Societal and Environmental wellbeing

Diversity, Non-Discrimination and Fairness

Transparency

Algorithmic accountability is the concept that organizations should be held responsible for the results of their programmed algorithms. The concept goes hand in hand with algorithmic transparency, which requires organizations be open about the purpose, structure and underlying actions of the algorithms used to search for, process and deliver information.

As the product of humans, algorithms can have issues resulting from human bias or simple oversight. Algorithmic accountability is promoted as a way to help such issues be recognized and corrected.

Trustworthy & Accountable AI: Goals

The Trustworthy & Accountable AI expert practice aims to:

1. Facilitate interactions between academia and industry
2. Disseminate and increase the impact of academic research
3. Educate digital accountancy & finance professionals

Collaboration opportunities can take different shapes:

- Data sharing and research collaboration
- Contract research/consulting
- Research funding (e.g., MSc theses and PhD projects)

Trustworthy & Accountable AI: People

Experts:

- Dr. Iuliana Sandu – Director
- Dr. Otto Koppius
- Prof. Marcel Rinsum
- Dr. Stephan Kramer
- Dr. Evelien Reusen
- Dr. Caspar David Peter
- Dr. Steven Vanhaverbeke

PhD candidates:

- Tamara Thuis
- Aljaž Sluga
- Sebastian Stirnkorb
- Albert Dongo
- Anoek Leonieke Holthuijsen

Trustworthy & Accountable AI: Research and educational interests

Erasmus Centre for **Data Analytics**

Trustworthy and accountable AI

- **The Audit of Algorithms**

- **Trust mechanisms in AI**

- **Types of accountability over AI**

- **Bias and incentives in AI development**

- **Transparency of AI**

- **Impact of AI on the accountancy profession**

- **Educating digital accountancy & financial professionals**

Trustworthy & Accountable AI: Relevant publications

- Sandu, M.I. & Koppius, O. (2019). Algorithms under control: An assertion-based framework for the audit of algorithms (working paper, November 2019).
- Dalla Via, N., Perego, P., & Van Rinsum, M. (2019). How accountability type influences information search processes and decision quality. *Accounting, Organizations and Society*, *75*, 79-91.
- Kramer, S., & Maas, V. S. (2016). Selective attention to performance measures and bias in subjective performance evaluations: an eye-tracking study. *Available at SSRN 2457941*.
- Stouthuysen, K., Teunis, I., Reusen, E., & Slabbinck, H. (2018). Initial trust and intentions to buy: The effect of vendor-specific guarantees, customer reviews and the role of online shopping experience☆. *Electronic Commerce Research and Applications*, *27*, 23-38.
- Vanhaverbeke, S., Balsmeier, B., & Doherr, T. (2019). *Corporate Financial Transparency and Credit Ratings*. Working paper.

Trustworthy & Accountable AI: Current Partners & Collaborators

Passion provides purpose, but data drives decisions

Erasmus Centre for Data Analytics

Dr. Iuliana Sandu

✉ sandu@rsm.nl

🌐 www.eur.nl/data

in https://www.linkedin.com/in/iuliana-sandu-021b633a/