

Explainability

Readings on what is meant by “explainability” (and related terms like “transparency” and “interpretability”) in data science and to what extent achieving explainability (or transparency or interpretability) in algorithms is morally important.

Explainability, Transparency, and Interpretability Readings

Fruit prices

Title	Citation
Transparency in Complex Computational Systems	(sep-privacy?)
How the Machine “Thinks”: Understanding Opacity in Machine Learning Algorithms	(sep-privacy?)
Algorithmic and Human Decision Making: For a Double Standard of Transparency	(sep-privacy?)
The Right to an Explanation	(sep-privacy?)
The Mythos of Model Interpretability	(sep-privacy?)

References