

Transparency of Artificial Intelligence

Regulatory Approaches and their Critiques

Transparency: A term that is both central and yet underdeveloped?

Since the early 2020s, laws relating to AI have been a major focus for the most important regulators globally. The most recent additions to the growing canon of AI regulation include the EU AI Act, an executive order issued by the US president and following efforts by federal agencies, multiple laws in China – particularly concerning generative AI – as well as the Framework Convention on AI by the Council of Europe. International coordination and cooperation in matters of AI regulation has also substantially increased, especially through established forums such as the OECD, the UN and the G7/G20. In almost all regulatory works to date, transparency has been pronounced as one of the most central and principles and ideas for effectively reigning in AI. However, so far, no common understanding of transparency in this context of what may be called “AI Law” has emerged. Adding further complications, terms like *Explainability*, *Interpretability* and *Understandability* have been introduced to the discourse around legal transparency – sometimes as synonyms, sometimes with distinctly different contents, sometimes even overlapping in meaning.

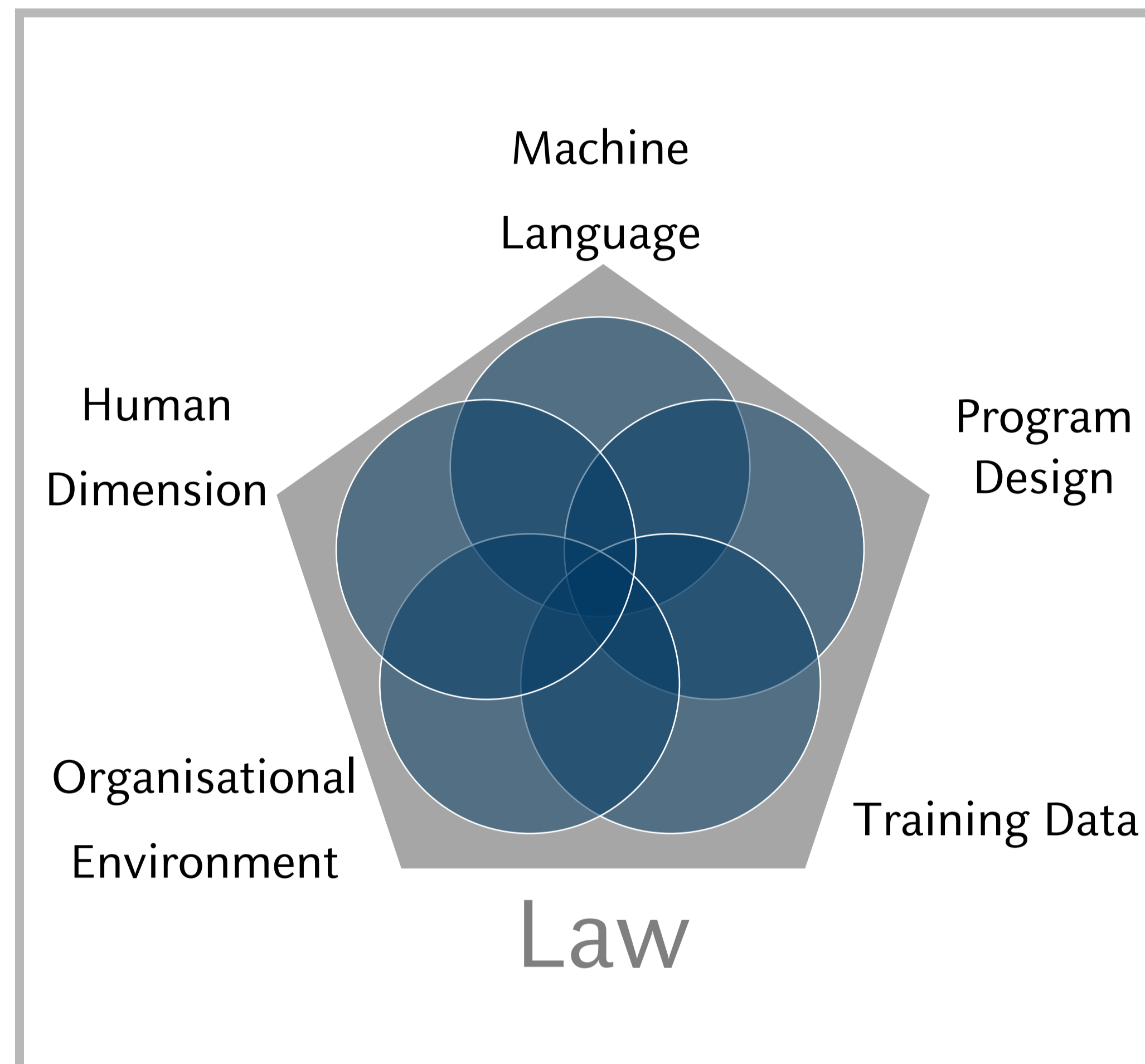
To aid in the operationalization of the actual legal norms based on the notion of transparency, a normative reflection on this concept under consideration of the multiple facets of AI Law affected and the ultimate goals of AI regulation in general is required – especially the idea of **effective remedy** for violation of rights through use of AI provides ample opportunity for inquiry.

An Instance of Legal AI Transparency

Art. 1.3 of the OECD Recommendation on Artificial Intelligence

AI Actors should commit to transparency and responsible disclosure regarding AI systems. To this end, they should provide meaningful information, appropriate to the context, and consistent with the state of art:

- i. to foster a general understanding of AI systems, including their capabilities and limitations,
- ii. to make stakeholders aware of their interactions with AI systems, including in the workplace,
- iii. where feasible and useful, to provide plain and easy-to-understand information on the sources of data/input, factors, processes and/or logic that led to the prediction, content, recommendation or decision, to enable those affected by an AI system to understand the output, and,
- iv. to provide information that enable those adversely affected by an AI system to challenge its output.



Opacity – Antagonist and Target

In many different regulatory acts and proposals as well as in lawmaking guidelines and international soft law, transparency is introduced as a concept to counter an explicitly or implicitly acknowledged opacity attributed to AI systems. At a first look, this makes sense – transparency is, after all, the definitional opposite of opacity. But seldom do regulators take the time to explore, in detail, what constitutes this opacity. The graphic to the left attempts a demonstration of the multi-factorial, „moving“ and interconnected factors constituting most understandings of AI opacity against the background of law, which is tasked with responding to the relevant concerns encompassed by opacity – either through operationalizing transparency in legal norms or by other regulatory means.



Unlike in the biblical tale of the Tower of Babel, depicted above in the rendition of Athanasius Kircher (1679), the cause for the multitude of understandings and „transparency-languages“ is not divine intervention, but rather i.a. **speed, cultural differences** between as well as initial **isolation** of regulatory efforts.



Legal Transparency beyond AI Regulation

In other areas of international law, transparency has been deployed and operationalized with different levels of success and vary foci. Among others, transparency can be found in rules relating to:

- Institutions (Inter- and Supranational)
- Governance (Democratic Processes and Lawmaking)
- Compliance (Environmental and Business Law)

These may offer structural references for future interpretation and implementation of AI-related transparency provisions – however, in the context of AI regulation, many transparency challenges are of technical or techno-social nature, which may require genuinely original analysis, thus presenting a frontier of legal transparency research.

Background and Effectiveness of Transparency

The main transparency *narratives* (1) revolve around procedural fairness, accountability, reduction of asymmetries, and public good – all of which are regularly noted as goals of AI regulation. However, transparency as a concept in law and corresponding narratives have been subject to scrutiny for many reasons, especially unclear effectiveness in supporting the aforementioned goals and influence on decisions.

Ideally, current regulatory approaches should take these critiques into account in order to ensure that transparency is successfully translated from a well-meaning principle into effective legal norms.