

Let's Talk AI with Barbara Steffen

Barbara Steffen¹

METAFrame Technologies GmbH,
barbara.steffen@metaframe.de

*"By far the greatest danger of artificial intelligence is that people conclude too early that they understand it."
- Eliezer Yudkowsky*

The Interviewer - Barbara Steffen



My Personal AI Mission:

Given the wide range of opinions about AI's potential and associated risks, I am motivated to help to structure this diversity and establish points of shared understanding.

Definitions and why they are Part of Interviewees' Profiles

Having conducted over 30 interviews on AI, I observed a wide range of opinions. This variance in experts' views may stem from their focus on different aspects of AI, their interpretations of key terms, or their perspectives on the consequences and mechanisms involved.

To provide clearer insights into these differences, I have decided to include profiles for all interviewees. These profiles offer a glimpse into their views and understanding of several key terms. This approach does not offer a uniform understanding but allows readers to interpret the experts' answers in the light of their personally intended meanings and definitions.

Why did I Choose these Terms for Definitions?

Artificial Intelligence: While everyone is familiar with AI, perceptions of what it entails can vary widely. In scientific research, a clear definition of AI is crucial for overall interpretation. Since there is no universally agreed upon definition, each interviewee was invited to provide their personal interpretation.

Trust: Trust has emerged as a central theme at AISoLA. Trust in a system or AI is a prerequisite for its use. Key questions include: When should we trust? When do we trust? How do we calibrate trust? It all comes back to, “What is trust? When is something trustworthy?” Opinions on this vary by person and discipline, but making these differences visible is the first step toward a more in-depth understanding of the concept.

Explainability: Explainability is a powerful lever for understanding and addressing AI. The term itself invites a wide range of opinions, influenced by the background and perspective of the expert. Clarifying what is meant by explainability and how to achieve it is critical.

Essential Elements of Human Capabilities: Especially when discussing Artificial General Intelligence (AGI), which aims to match human intelligence on all fronts, it is important to first understand and ideally agree on what human-level intelligence and capabilities are. This understanding allows us to measure and evaluate if and when AI reaches or exceeds it.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

