Let's Talk AI with Georg Borges

Georg Borges[1] and Barbara Steffen[2]

[1] Saarland University, Faculty of Law, Germany,
[2] ls.borges@uni-saarland.de
[3] METAFrame Technologies GmbH,
barbara.steffen@metaframe.de

"Frankenstein is alive - AI will transform artefacts into actors that are capable of self-chosen actions and able to interact independently with humans. In view of this development, we need to develop a legal framework that takes these new capabilities into account, particularly one that defines the space and boundaries for human-machine relationships. "

The Interviewee - Georg Borges

My Personal AI Mission:
To see the creation of a new legal framework for AI systems.

My Takes on AI

Artificial Intelligence: Methods to enable machines to perform tasks similar to human thinking.

Trust: Accepting (without further precaution) a risk regarding a future event or action, in particular future behaviour of another person.

Explainability: I distinguish explainability from justifiability. Justifiability relates to the question if a decision can be considered a convincing conclusion based on arguments whereas explainability relates to the question how a specific decision was actually reached.

Essential Elements of Human Capabilities: self-awareness; empathy; communication.

The Interview

Barbara *Hello and welcome to Professor Georg Borges. Thank you for your time. Could you please briefly introduce yourself and your personal relationship to artificial intelligence?*

Georg Thank you for the opportunity. I am Georg Borges, a professor of law at Saarland University, Germany, and the director of the Institute of Legal Informatics at the same university. I have worked in the field of IT law for over twenty years and have been engaged in the intersection of AI and law for about nine years. I still recall my first presentation on the legal framework for AI in 2014 when robots with high communicative potential started to emerge, and we questioned their potential role in society.

"I also believe it is possible to fall in love with an AI system, so there is no limit to this."

Barbara *Which specific AI challenges does your research currently address?*

Georg My research currently addresses a multitude of challenges related to almost every aspect of law and AI. Our law is human-centric; we address humans, not machines. The most significant and far-reaching question that has arisen from this is the question of e-personhood [1, 3, 5, 6]: should we consider machines as legal subjects or merely as objects? This distinction is critical in private law [4]. If we choose to accept e-persons, then your computer might be your boss or own the shares of your company. Hence, determining the status of an AI system is of utmost importance for almost everything. It permeates every aspect of the law where so we must decide whether a norm can address a machine or not, or if an AI system should be considered a machine or something different, something new.

"Some of my Japanese colleagues dream of a society where humans and AI interact as equals. That might be the future."

Barbara *Does this mean that we might need a completely different legal framework? Or is the current framework sufficient, covering all the essential aspects, and only needs to be adapted or extended to better meet the new challenges?*

Georg I believe we need a completely new framework. However, the question is, what constitutes "completely new"? When the Internet emerged, at first many people said there was not much novelty. Yet, it changed society and the law significantly. If we were to start addressing machines as we currently address only fellow human beings– and I think, we will, that would also lead to a massive shift. Developing norms to address machines is a significant challenge [2]. So, if we start addressing machines, we must reconsider the content of the norms, how we construct norms, and our standard of expectation. The way we develop and promulgate these new norms will differ from the one we have been employing for norms addressing human beings. Whether you call it law or technical norms as

an appendix to legal norms, I do not mind, as long as the effect is a framework of normative requirements that address machines. This is quite a novel concept.

Barbara *What role should law or regulation play in the adoption of AI to ensure or guarantee trust in AI? Should the focus be on trust in the machine, trust in its functionality, or trust in the organization providing the product? Or is it about sufficient and trustworthy oversight, which means that we outsource the trust issue to regulation, which sets and controls the conditions that are worthy of the user's trust?*

Georg Trust in AI is a complex concept [7, 9]. We have learned a lot in this conference about what trust in AI could mean. There are different interpretations, even among lawyers. So, it is challenging to answer such a question. For instance, in one presentation, the authors wanted to determine whether a seal or certificate on certain aspects of trustworthiness would influence consumers' trust in the AI system. They found out it did not make any difference. So, what forms the foundation of our trust? Is it the reputation of the technology's producer? The service provider's reputation? The legal framework guaranteed by state power that is protecting individuals? Or is it a combination of all these factors? Likely, several elements together constitute a basis for trust. One of the most interesting aspects in this field of research will be to identify these different factors of trust and try to establish a solid foundation of trust by strengthening these factors. But then, we must also consider why we need this trust. What role does it play in AI? Probably the willingness to use it. Yet I use things I do not trust entirely. Moreover, I even lack a notable amount of trust in many of the internet services or products I use daily, or in many of the people I talk to. However, I have a high level of trust in my family, friends, partners, and close staff. Likewise, we have several levels of trust relevant to different areas of law, and the foundation of that trust is quite different in each case. When we talk about trust, which I have been working on for several years, we need to distinguish what we are referring to. Then we can establish the factors for trust and work on those factors. I agree that trust as a concept is fundamental to human interaction and the use of technology. But when we try to support trustworthiness, we become aware of the complexity of trust itself.

" If we were to start addressing machines as we currently address only fellow human beings – and I think, we will, that would also lead to a massive shift. . . . . we would need a completely new [legal] framework"

Barbara *There's now a company called Inflection AI that's working on Pi, which stands for personal intelligence. Pi is basically designed to give you the experience of chatting with your own personal assistant. Inflection AI is marketing Pi as a personal chief of staff that is always in your pocket, always available to you. You can interact with it whenever you want, and it can also take over mundane and boring tasks like managing your schedule or reaching out to people with standardized emails and stuff like that. So, I wonder if we start interacting more*

and more with these kinds of AI-powered applications, if it's not natural to start trusting it more and more over time. It gives great, helpful, personalized answers, and it engages in long and extensive conversations with you. People may start to have trouble really distinguishing between machine and human interaction. They know it's 'just' a machine that responds in a very human way, but it feels more like chatting with a human, but it's not a human. So how much trust should we have and how can we make sure that people stay skeptical or suspicious enough to not lose all distance after the 10th time they use it? It gets normal, right?

Georg This question has been explored in many movies and novels. Personally, I would like to possess such an assistant, as depicted in several films. I also believe it is possible to fall in love with an AI system, so there is no limit to interaction with machines. However, from a legal perspective, the question will arise: if Humans can enter into emotional relationships with machines, what should be the legal status of such relationships? Can I marry my robot? – Certainly not, as long as the robot is not considered a legal entity, but in any case, the interaction will at least be protected by the personality rights of the human. Emotional relationships between humans and AI systems will create specific risks because such an app can be controlled by a company or many apps can collect data from many use cases, making it very powerful. Therefore, we certainly need a robust legal framework to minimize the risk of misuse of AI assistants. That is why I am here, to build up such a legal framework that allows us to use such powerful AI systems, which we will undoubtedly do some day. The idea of having such an assistant, which I cannot see myself being able to afford on my income, is appealing. If they can do the dishwashing, even better.

"Our law is human-centric; we address humans, not machines. The most significant and far-reaching question is that of e-personhood: should we consider machines as legal subjects or merely as objects?"

Barbara *In terms of future technical capabilities of artificial intelligence, on a scale of 1 to 10, where 1 is the artificial intelligence systems we see today, like ChatGPT, which are specific to certain tasks, and 10 is something like general artificial intelligence systems that potentially surpass human capabilities. What should we expect?*

Georg I am not a technician, so I have limited insight on what will happen in the technical field. But from a layperson's perspective, my guess would be 8 or 9. I think progress will continue, and the advancement we have seen so far is already very interesting. By combining several tools, it should be possible to create a very powerful tool that appears superintelligent. My guess is that it would not be one general AI, but a combination of different systems will appear like a general AI. That is probably possible.

Barbara *Given these possible outlooks, how do you personally feel about the much-discussed spectrum from utopia to dystopia? What can we expect?*

Georg Predicting the future is challenging. I am an optimist by nature, so I hope for the best. I think there is a fair chance to avoid a dystopian future. Since the mid-60s, mankind has been possessing the potential and means to destroy significant parts of the world, to kill more than a billion people, and to make this planet uninhabitable. But we have not done such a thing so far. We have been smart enough not to use nuclear weapons all at once, and we have survived. That gives me some hope that we can avoid making foolish mistakes in the future, such as using AI in a way that would ultimately destroy us.

Barbara *Reflecting on the last few days and the presentations you heard from other disciplines, what was the most interesting insight you got?*

Georg That is difficult to say. There were many interesting insights. Surprisingly, I found myself most challenged by something I took for granted. In one empirical study on the effect of certificates on the trustworthiness of AI systems [8], the authors found that such seals did not have any measurable impact on the trust of people in AI systems. This finding contradicts a strong belief I held, which I now have to reevaluate. That was incredibly inspiring because when you take something for granted and build your concepts on it, and it turns out to be incorrect, you should modify your idea. Interestingly, these challenges did not come from legal presentations but from other disciplines.

"I agree that trust as a concept is fundamental to human interaction and the use of technology. But when we try to support trustworthiness, we become aware of the complexity of trust itself."

Barbara *Is there a particular research question you would like to see addressed from an interdisciplinary perspective?*

Georg Not one in particular, I believe most research questions would benefit from being addressed from an interdisciplinary perspective.

Barbara *From your personal perspective, what should be our AI vision?*

Georg Our AI vision? Some of my Japanese colleagues dream of a society where humans and AI interact as equals. That might be the future. It would be challenging. I do not know if it is a good vision or not, but certainly, using AI as a trusted friend of humans would be my most optimistic vision.

Barbara *Perfect. Do you have anything to add?*

Georg I'm afraid not. Your questions were incredibly interesting. and I hope I provided satisfactory answers. Thank you so much for engaging with me.

Barbara *Definitely, your answers are very informative and complement the other interviews. Thank you very much, Georg, for your time and insights on AI from a legal perspective. Have a great time at AISoLA!*

Georg Thank you very much. Goodbye.

References

1. Borges, Georg, A Legal Framework for Autonomous Systems, in: Borges, Georg/Sorge, Christoph (eds.), Law and Technology in a Global Digital Society: Autonomous Systems, Big Data, IT Security and Legal Tech, Springer, 2022, p. 3, 5 f.
2. Borges, Georg, Liability for AI Systems Under Current and Future Law: An Overview of the Key Changes Envisioned by the Proposal of an EU-Directive on Liability for AI, 24(1) Cri, 2023, p. 4, para. 21 f.
3. Dremliuga, Roman, Kuznetcov, Pavel, & Mamychev, Alexey, Criteria for Recognition of AI as a Legal Person, 12(3) JPL, 2019, pp. 105-112.
4. Kornilakis, Angelos, Legal Personhood for Autonomous AI: Practical Consequences in Private Law, in: Kornilakis, Angelos/Nouskalis, Georgios/ Pergantis, Vassilis/Tzimas, Themistoklis (eds.), Artificial Intelligence and Normative Challenges: International and Comparative Legal Perspectives, Springer, 2023, pp. 27-37.
5. Laukyte, Migle, AI as a Legal Person, 19 ICAIL, 2019, pp. 201-213.
6. Militsyna, Kateryna, Legal Personhood for Artificial Intelligence: Pro, Contra, Abstain? 122 Teise VUP, 2022, pp. 150-158.
7. Razin, Yosef S./Alexander, Kristen, Developing AI Trust: From Theory to Testing and the Myths in Between, ITEA 2024, available at: `https://itea.org/journals/volume-45-1/` (last accessed 21 May 2024).
8. Wischnewski, Magdalena Wischnewski / Nicole Krämer, Nicole /Janiesch, Christian Janiesch / Müller, Emmanual Müller / Theodor Schnitzler, Theodor /Neven, Carina Neven,: In Seal We Trust? Investigating the Effect of Certifications on Perceived Trustworthiness of AI Systems, Human-Machine Communication, 2023.
9. Yang, Rongbin/Wibowo, Santoso, User Trust in Artificial Intelligence: A Comprehensive Conceptual Framework, 32(4) Electronic Markets, 2022, p. 2053, 2056 ff.