Let's Talk AI with Martin Leucker

Martin Leucker[1] and Barbara Steffen[2]

[1] University of Lübeck, Institute for Software Engineering and Programming Languages, Germany,
leucker@isp.uni-luebeck.de
[2] METAFrame Technologies GmbH,
barbara.steffen@metaframe.de

"Analysis and verification methods must be developed if AI is used in safety critical domains."

The Interviewee - Martin Leucker

My Personal AI Mission:
is to make AI techniques more trustworthy and create further methods and methodology for designing and building powerful and safe AI-based solutions efficiently: SE4AI.

My Takes on AI

Artificial Intelligence: A powerful tool that may be applied in many areas by assisting in daily work activities.

Trust: In an AI-based system trust is the mental attitude of a human that the system is doing what it is supposed to do in given circumstances. It involves correctness of the system but many more attributes like privacy etc.

Explainability: For AI-based systems explainability is their capability of illustrating in a human understandable way how it works.

Essential Elements of Human Capabilities: Humans are capable of using context information much better in their daily reasoning than artificial systems.

The Interview

Barbara *Today I have the pleasure of speaking with Professor Martin Leucker. Please introduce yourself and your relationship to artificial intelligence.*

Martin Thanks, Barbara, for giving me the possibility to give this interview with you. My name is Martin Leucker and I'm heading the Institute for Software Engineering and Programming Languages at the University of Lübeck. At the same time, I'm also CEO of the UniTransfer-Klinik Lübeck GmbH, which is a technology transfer company that is mostly owned by the University of Lübeck, as well as the UKSH, which is the university hospital in Schleswig-Holstein. As such, I'm developing software or I'm working on methods for developing software, so I'm working on software engineering techniques with lots of applications in the medical domain. Traditional software engineering techniques now have to be adapted to be ready for AI technology, and that is basically what we do research on.

"If you take medical devices, there are strict rules under which medical devices can get certified and put into the market. Now, these kinds of rules have to be adapted to cater for AI components."

Barbara *Do you have examples of one or two research questions that specifically address AI?*

Martin For example, if you take medical devices, there are strict rules under which medical devices can get certified and put into the market [8]. Now, these kinds of rules have to be adapted to cater for AI components. So, when now AI components come into medical devices, you have to have new techniques for certifying such devices, and that is one of the research questions.

Barbara *Could you elaborate on that?*

Martin For example, one of the rules says that the software has to be intensively tested before you put it into the market, which makes sense, of course, as these are safety-critical systems. Now, what does it mean? If we learn a control strategy, for example, for a medical device, given certain kind of data, then we have to test it differently than when programming the control manually.

At the same time, there are lots of techniques where you continuously learn. You observe what the system is doing in practice and the software changes dynamically. So far, such kind of AI technology is not allowed to put into the market, because according to current rules, once you test the system, it is not allowed to change [3]. So, one of the real challenges in this domain is, could we eventually, sooner or later, also allow software that is changing itself, learning in practice, for allowing it to the market? But that, of course, would require that we have safeguards aside that kind of make sure that whatever the system learns is still within some kind of safety boundaries [6].

Barbara *Building on that, what role does trust play in AI adoption?*

Martin Trust is a very important concept, but at the same time, I think trust is much more than correctness. And in many cases, it's misunderstood. Trust does not only mean that the system is working correctly. But we have to guarantee many different forms of trust. For example, when a doctor uses a device, it's important that the device works correctly from a provable, mathematical perspective, but it's also important that the doctor gets the feeling, gets the understanding that the device is working correctly. At the same time, for example, many machine learning technologies are based on data, and that means that the patient should give data to the companies or to the research institutions to allow them to build new solutions based on the data. Of course, privacy issues, anonymization issues, and so on, they have to be respected. But even if you would, on a technical level, be able to guarantee such kind of privacy issues, it also means that the people have to trust the solution. So, it's a mixture of technical solutions, but also has lots of social aspects as well [4].

"For me, the role of AI is to be a smart assistant, but it always means that we have a human supervisor [...]."

Barbara *Do you have key measures in mind that need to be in place to ensure ethical AI adoption?*

Martin Well, regarding ethical AI, I think it's important that on one hand, we understand, we learn, we come up with rules what is ethical in this context. But at the same time, it's important that these rules materialize in law. Ethical aspects are fine, but there also has to be a punishment if you don't follow them. I think it's important that, let's say, philosophers, people working on ethics work in AI. But it's also very important that we collaborate with people from law that also tell us in which way such kind of ethical implications will eventually be enforceable.

Barbara *In terms of the future technical capabilities of AI, on a scale of 1 to 10, what do you think will be possible? One being the artificial intelligence tools we see today, like ChatGPT, and ten being artificial general intelligence that surpasses human capabilities?*

Martin Personally, put it in numbers, maybe it would be like 7. I think that the current state of the art will advance a lot. But as you see from ChatGPT, for example, it looks quite nice, but when you really dig into it, then you see that a lot of things are not really well understood. So, for me, the role of AI is to be a smart assistant, but it always means that we have a human supervisor that is in the end in control of when to apply AI-based solutions or when to decide certain kinds of things. To give an example for the medical domain, it will always be the doctor that has the final word. Maybe he or she has a lot of assistants that are based on AI, that find certain kind of malicious cells in the radiology picture and so on. But the final assessment has to be done by a human. Likewise, an AI-assistant may suggest a treatment based on what has worked in the past, but the choice of treatment has to be under the supervision of a human.

Barbara *There has been a lot of discussion about possible future scenarios for artificial intelligence, from dystopian to utopian. Where would you position yourself based on your understanding of AI's future capabilities, etc.?*

Martin I think that the sensible application of AI techniques is in general a good idea. I'm completely sure AI will not take over the world [2]. All these kinds of horror scenarios that are often discussed in the media, I think it's completely over-exaggerated. I think that one can make a very positive use of AI solutions in the next years. Nevertheless, it is good to have these discussions and to constantly monitor and ensure that AI is developed in the right direction, for the benefit of our society.

Barbara *Reflecting on the past few days, what insight was particularly interesting to you?*

Martin Well, here at AISoLA there were many different tracks. I think it was important to have tracks on explainable AI, to more understand notions of trustworthiness where philosophers were explaining their views, where at the same time also lawyers were explaining their views. That I liked a lot. Then there was a track on digital humanism which is I think also very important to bridge the gap between, let's say, computer science, AI and humanities. Then I also liked a lot the track on future programming based on AI-based tools, because that fits my software engineering background. Of course, I also like the healthcare track, as it shows the applications in the medical domain. That is very important for my work, too [7].

"[Fake news] always [have] been around, but it's now so easy to create fake news in a massive manner."

Barbara *You yourself are used to working in an interdisciplinary environment, right?*

Martin Yes. We have lots of projects with the application in the medical domain, so we have to work with doctors. But at the same time, we also have to follow the norms and the rules. We have to understand them, maybe also to criticize them and say where to extend them. That has something to do with reading laws and norms and so on, which we were not educated for. So it's good to cooperate with lawyers in these aspects. At the same time, also ethical aspects play a role in our work, again a direction for interdisciplinary cooperation [5].

Barbara *Are there any lessons learned from these collaborations that you would like to see applied to artificial intelligence and its evolution and adoption? Maybe things that could be done similarly or things that didn't go so well.*

Martin I would say that we are on a very good track on one hand, but also on a slow track. One of the difficulties is, of course, that every discipline speaks a different language. It just takes a while to understand the other groups, to understand their thinking, their terms, their definitions. I think we are on a good track, but of course, there is still a lot of work, a lot of research to be

done. We're concentrating on the medical domain, but then at the same time, the automotive industry is very important, autonomous driving. There's also a huge area where certain norms play a very big role. We haven't compared the different approaches yet. In some sense, right now we go into limited application areas, but you also have to harmonize between different areas. What is common in the different application areas? What is different?

Barbara *Is there a particular research question or topic that you would like to see explored in more depth from an interdisciplinary perspective?*

Martin From an interdisciplinary perspective, I think it's very important that people from the legal domain cooperate with computer scientists. And also, people from normative institutes like ISO, DIN, and so on, that they participate in interdisciplinary research. At the same time, for my specific personal research area, the verification of AI-based solutions is an extremely important topic that we have to study in more depth. Right now, many people come up with AI-based solutions, but they don't really have verification techniques. And verification is kind of the precursor to get systems approved by authorities.

Barbara *Is there already a framework in place that provides structure and clarifies the expectations and roles of the various disciplines involved? Or is this something we need to build?*

Martin I think we're currently building it. I mean, it's events like AISoLA, for example. They are a perfect opportunity where different kinds of communities can meet, can exchange ideas, where we can learn from the other groups. We have interdisciplinary projects where we also, of course, have to collaborate. But I think we're on a good track, but it's still a long way to go.

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

Martin Well, to me, it looks like AI is going to be an important tool in building a digitalized world. Meaning building new systems where we have AI-based components, where we have a lot of interaction with the help of AI-based systems. But at the same time, AI is a tool and not more. So I don't believe in, as I mentioned before, AI will take over the world or something like that. It's supposed to be a tool, in my understanding, it will be a powerful tool, but also not more than a tool.

Barbara *Do you see a potential for increased misuse?*

Martin Of course. Like every new technology, it can be misused. We have the problem of fake news that are easily created now using ChatGPT. You can get very nice texts that look reasonable but are completely fake. The same with pictures and videos and so on. It's definitely a threat to society, which has to be solved by technology that, for example, identifies news as fake news [1]. But at the same time, and may be even more important, it has to be solved sociological, for example, train people to really look for the references and to understand what is real information, what is fake information [9]. In a sense, it's nothing really new. It always has been around, but it's now so easy to create fake news in a

massive manner. And that means also that the people have to be much more careful than before.

Barbara *Do you think users are sufficiently aware of this? Or do we need to teach them to be more skeptical in their interactions*

Martin That's a hard question. I think it's a little bit of both. On one hand, people are aware of it in a sense that some people are very critical about AI. At the same time, they send their data to companies like Meta via Instagram and so on. So, to my feeling, people are aware of problems, but they don't understand the technology to the extent that would help them to really develop the right understanding and right treatment of AI solutions. So, in some sense, I think we have to train people in the way that they can judge what is a reasonable approach, when to share my data and when not to share my data. So right now, it's a little bit blurry that some people are frightened, but they don't know why. They don't know when to be frightened and that results sometimes in irrational behavior.

"Right now, many people come up with AI- based solutions, but they don't really have verification techniques [...]. And verification is kind of the precursor to get systems approved by authorities."

Barbara *Is part of the challenge that organizations are not sufficiently transparent about what they do with this data?*

Martin Well, that's one aspect. However, it is not only the organizations but also the users of technology: I think for most technology, even if you are a person that is not really working on this technology but only with this technology, you still have to understand it to a certain extent. For example, if you drive a car, you have to have a limited knowledge how a car works to be able to understand what is a dangerous situation, what is not so dangerous and so on. And this training of people is missing when it comes to AI. So, we need a general understanding in society about AI, what is possible, what is not possible, when is some fear rational and when is it irrational.

Barbara *Is there anything else you would like to add?*

Martin Not really.

Barbara *Then thank you Martin for your time and your perspectives on AI, especially from the healthcare perspective. Have a great evening!*

Martin Thanks a lot. And thanks for taking this initiative!

References

1. Islam, M.R., Liu, S., Wang, X. et al. Deep learning for misinformation detection on online social networks: a survey and new perspectives. Soc. Netw. Anal. Min. 10, 82 (2020). https://doi.org/10.1007/s13278-020-00696-x

2. Lewis, Tanya (2015-01-12). "Don't Let Artificial Intelligence Take Over, Top Scientists Warn". LiveScience. Purch. https://www.livescience.com/49419-artificial-intelligence-dangers-letter.html, Retrieved March 17, 2024.
3. Ludwig Pechmann, Martin Mildner, Tim Suthau, Martin Leucker: Regulatorische Anforderungen an Lösungen der künstlichen Intelligenz im Gesundheitswesen.Künstliche Intelligenz im Gesundheitswesen 2022: 175-198
4. Ethics guidelines for trustworthy AI, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
5. Project pages of the KI-SIGS projects for a typical project and the interplay of different disciplines when building AI-based medical devices, https://ki-sigs.de
6. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback, https://www.fda.gov/files/medical\%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf
7. https://2023-aisola.isola-conference.org
8. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX\%3A32017R0745
9. https://www.edps.europa.eu/press-publications/publications/techsonar/fake-news-detection_en