Reflections on AI
Q&A with
Prof. Emma Ruttkamp-Bloem

"AI technology has, from its inception, been meant to augment what it means to be human..."

The TUM IEAI had the pleasure of speaking with Emma Ruttkamp-Bloem, Professor and Head of the Department of Philosophy, University of Pretoria, South Africa.

Q: What is the biggest misconception about AI?

A: I think there are a lot of people who think that we are already at human level AI or that we are very close to that and that there is some kind of compulsory pressure on everybody to engage with AI technology, especially companies in the private sector and also in academia, and that is exceptionally dangerous.

I'll give you one example of harm that could come from this current misconception. I hear lots of people saying that generative AI applications would be able to change education, specifically in Africa, forever - for the good. My first reaction is yes, it will change it, but not necessarily for the good. We all know that it is not necessarily the case. Firstly, people who live in rural areas in Africa will not necessarily have access to these kinds of apps, no matter what these apps supposedly can offer them. But secondly, we know about all the problems around the quality of the responses generated by these apps and if specifically in Africa, where children might not have access to teachers who can guide their engagement with these kinds of apps, that is really, really, really dangerous.

Q: What is the most important question in AI ethics right now?

A: The whole debate between AI ethics on the one hand, and international human rights law on the other hand, depends a lot on the context and the platform. In some cases AI ethics can't do for governance what human

rights law can do. It is soft law. If we think in terms of the UN advisory body (that I'm a member of), we make it very clear as our fifth principle in our interim report that came out in December (2023) that international laws are the bottom line for any kind of governance of AI. But what is forgotten? Or overlooked?

AI ethics is not, at least as far as I'm concerned, simply unapplied ethics. And it is not about top down sets of principles and values that hang somehow in the air that nobody cares to actualize or to explain. We absolutely don't need that kind of AI ethics.

For me, AI ethics is the enabler of responsible AI and the first kind of flag for human rights digressions, and it is not a threat to human rights. It is absolutely a tool. And in the UNESCO recommendation we also defined AI ethics as a dynamic system. It is about reasoning. It's almost a meta ethics. It is about sitting and reflecting about where is it that our norms and values come from.

Q: What is the role of academia when it comes to the regulation of AI?

A: For me, the most important thing is that the voices from academia should make clear and illustrate how important transdisciplinary and interdisciplinary research is to inform governance. How I see our role is to help unpack racial discrimination in answers that are generated by generative AI to teach our students critical AI skills to explain why it is needed and what those things should be. And from the tech side, to show how important it is to interact with the humanities, for instance, and the law in order to be able to unpack these kinds of issues, and to understand why it is necessary for us to have these conversations.

It is really unfair on the one hand, and it's also irresponsible to expect of the tech community

Reflections on AI
Q&A with
Prof. Emma Ruttkamp-Bloem

"AI technology has, from its inception, been meant to augment what it means to be human..."

to be able to answer questions around structural buyers, how the AI business model in fact supports current geopolitical power play globally.

For those kinds of conversations, we need to have transdisciplinary conversations and the humanities need to understand how the technology works. It is for us to put the proof in the pudding to show what kind of research can really come out of these kinds of research and show what it can do for governance.

Q: In what ways do we need to understand concepts of *responsibility* and *agency* to inform the development and use of AI?

A: I say a lot about the why of AI ethics. In my work, the 'what' is the principles and values and the 'how' is the responses from the tech world to test an algorithm for buyers, or for transparency, of expandability or whatever the case might be. But I'm a philosopher, so I really think it is important for us to also understand the 'why'.

For me, the why relates to agency. It relates to ensuring that AI technology is used to support and enhance human agency and human agency for what? I think that in the case of AI ethics, it is not political entitlements, but ethical entitlements that place positive duties on governments, and on all AI actors, not just governments, to ensure that AI in the development, in the employment and use of AI technology do not in some way threaten the ability every citizen has to attain a life of well-being.

The other thing is this notion of what it means to be human. That relates directly to the right to integrity, the right not to be manipulated, the right to be the person that you want to be, the person that you find yourself to be. AI technology has, from its inception, been meant

to augment what it means to be human in ways that will not harm the core of what it means to be human.

Q: How and why should we bring together *global* and *local* perspectives on AI governance?

A: If we think, for instance, of the African continent, which is my continent, we have the youngest median age in the world. AI is developing very fast. I also always say a lot about how intergenerational justice is important to keep in mind when we think of AI governance and policy making, because we are changing the world that future generations will be living in and making decisions on how that world will be regulated. And that means that the whole world needs to be part of this conversation.

Secondly, if we talk about leaving no one behind. We must be very sure that we mean no one in the North and no one in the South. Usually when people say that they think 'oh the poor people in the Global South'. In fact, the whole world is at risk of harm from AI technology and if there is one pocket (and that's the third reason), if there is one place where there is no protection, it is problematic. In some cases, international global regulation is all that some people will have to protect them.

It is really important to understand that there are different articulations of shared values. We shouldn't only talk in terms of cultural diversity to the point that we don't understand each other anymore. We should understand and be open to unravel how different cultures unpack the same value, for instance, the value of privacy. It will be differently unpacked from a collectivist than in an individualist ethical tradition. But it does not mean that it's not respected and that it's not taken seriously and people don't understand what it means. We

"AI technology has, from its inception, been meant to augment what it means to be human..."

need to have sensitivity, so that we protect everyone.

Q: We often say that AI is changing or transforming the world. To what extent is AI changing us as humans?

A: It challenges us, perhaps in ways that we have never been challenged, because it asks us to be very honest about what it means to be human, and to be very honest about why it is worth protecting what it means to be human. And those are very hard questions to respond to, especially as I said, if you look around you in the world. So we need also to speed up conversations, for instance, in disciplines like philosophy, that are 3000 years old, on what it means to be human and to really pull up our socks.

Meet the Expert:

Emma Ruttkamp-Bloem, PhD, is a philosopher of science and technology, an AI ethics policy advisor, and a machine ethics researcher. Emma is a member of the newly convened UN Secretary General's AI Advisory Body. She is also the current Chairperson of the UNESCO World Commission on the Ethics of Scientific Knowledge and Technology (COMEST). She is also the South Africa representative for the Responsible AI Network – Africa (RAIN-Africa), co-founded with the IEAI. Currently, she is the Head of the Department of Philosophy, University of Pretoria, South Africa.

Disclaimer: Please note that the text version of this interview has undergone slight edits for clarity and conciseness.