Reflections on AI

Q&A with

Rafael A. Calvo

"The most difficult problem we face today, is how AI can support human autonomy."

The TUM IEAI had the pleasure of speaking with Professor Rafael A. Calvo, Professor and Director of Research at the Dyson School of Design Engineering at the Imperial College London, on the topic of "Value Tensions in the Design Engineering of Intelligent Systems".

Q: What is the biggest misconception about Artificial Intelligence?

I believe the biggest misconception conception is about human autonomy, we worry a lot about machine autonomy, but the most difficult problem we face today is how AI can support human autonomy.

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

The most important question is how to support human autonomy. This is very difficult because the concept of autonomy has been very diverse both in philosophy and psychology and in technology, I think technology design has the possibility and the ability to bring a lot of these concepts back to something that we can use in the design and evaluation of intelligent technologies.

What are some of the key ethical dilemmas that engineers encounter when working with AI, and how does this shape their decisions?

I think the most difficult dilemmas that engineers and designers face are the tensions they often find between the benefits to the company they work for and the impact that the technologies will have on the other stakeholders in the general public and society as a whole. They have a motivation that is to drive business and drive a technology that is aligned with those business goals, but then they have these other motivations about being responsible and ethical, and this is a very common tension.

What is the role of academia when it comes to the role of trustworthy and responsible AI?

I think academia can play the role of an external referee or mediator between the wider public, the media, and policymakers who do not necessarily understand how AI works and the businesses that have these tensions and conflicts of interest. So academics external to those technology companies can advise the general public or the legislators on how to control these in ways that do not destroy innovation but, on the other hand, remain responsible and support societal well-being.

How and why should engineers question the values by which they evaluate the systems they create to promote responsibility in development and implementation?

Each individual engineer comes to a project with their own personal background and experiences, and these experiences will shape the way they see the world, so engineers and designers will shape technologies in the ways they see the world improving. So it is important to understand and reflect on our values so we can design technologies aligned with the values of the general population, not just our own. This is quite important because we have to understand the impact that we will have on general society, which is probably why

diversity in engineering and design teams is so important. It is the responsibility of product managers and directors of a company to make sure that these teams are not just multidisciplinary but also multicultural.

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

The impact of technology goes way beyond the provision of material benefits. It will support or hinder the possibility of fulfilling our most basic psychological needs. Technologies can support our sense of autonomy, competence, and relatedness. When they support these psychological needs, they will allow us to be happier and more motivated in engaging in activities that the technology mediates.

However, technologies can also hinder our ability to satisfy those psychological needs. They can feel controlled; when they feel controlled, we are less likely to engage or be motivated to do that activity. They can feel as limiting our sense of competence, of stability to fulfill that particular behavior or task, and therefore, we will feel disengaged. They can also help or hinder our need to belong to a community; we can see this with social media, for example: rather than connecting us to others, it can make us feel more isolated.

So technologies can support or hinder our psychological needs, and this has an impact on the way we are as human beings; if we are not motivated to engage in certain behaviors, this could lower our chances of taking care of our health, we could become less motivated to study or to be curious to learn new things and

these behaviors happen when we are intrinsically motivated to do so. Thus, technologies can change these intrinsic motivations and therefore change our essential human nature great as it is.

Meet the Expert:

[Rafael A. Calvo](#) is the Professor and Director of Research at the Dyson School of Design Engineering, Imperial College London. He is also co-lead at the Leverhulme Centre for the Future of Intelligence and co-editor of the IEEE Transactions on Technology and Society. He is the director of the Wellbeing Technology Lab, that focuses on the design of systems that support wellbeing in mental health, medicine, and education. He is a keen player in the push for more ethical technology design, is the recipient of five teaching awards, and published four books and over 200 articles in fields of HCI, wellbeing-supportive design, learning technologies, affective computing, and computational intelligence.