

Chapter 7

Ethics of AI

ABSTRACT

While the ‘AI ethics’ debate has been raging for some time, there are legitimate questions about employing AI and automation that are explored in this chapter. AI ethics isn’t just about doing the right thing or making the best AI systems possible, it’s also about who wields power and how AI affects the balance of power in everything it touches across society including businesses, institutions, and governments – ultimately, affecting the privacy and human rights of all individuals globally. AI must be developed with an understanding of who humans are collectively and in groups (anthropology and sociology), as well as who we are individually (psychology), and how our individual brains work (cognitive science), in tandem with current thinking on global cultural ethics and corresponding philosophies and laws. One thing is clear, without ethics, all visions of AI’s future will have little impact.

“It’s not artificial intelligence I’m worried about, it’s human stupidity.” –Neil Jacobstein, Chair of the AI and Robotics track at Singularity University

INTRODUCTION

Artificial intelligence (AI) is a key technology that will profoundly change our economic, political and social structures in the future. Equally important is a

DOI: 10.4018/978-1-7998-4607-9.ch007

widely held belief that AI represents one of the great human rights challenges of the 21st century (Elsayed–Ali, 2019). Although the *AI ethics* debate has been raging for some time, there are legitimate questions about employing AI and automation. It is not just about doing the right thing or making the best AI systems possible, it is also about who wields power and how AI affects the balance of power in everything it touches across societies, businesses, institutions and governments. Ultimately, AI affects the privacy and human rights of all individuals globally.

It is important to pause here and reflect on our individuality: humans across the globe are *different*. Hence, AI algorithms and automated processes will not work equally effectively worldwide, because different regions have different cultural models of what constitutes sociability and ethics. The values, morals and ethical choices that each individual makes on a daily basis are rooted in their cultural framework and sense of sociability. On their own, cultures, legal standards and regulatory frameworks will not get us out of the predicament we find ourselves in regarding AI. For that, we not only need tech developers to read some moral philosophy, we also need politicians, business leaders, community activists and citizens to do the same. Presently, we appear to be dancing around the edges of the issue. We treat tech design and development as if it is inevitable. Thus, we reactively minimise risks after the fact rather than look more deeply at the values, goals and moral commitments being built into the technologies.

AI must be developed with an understanding of who humans are collectively and in groups (e.g. anthropology and sociology), who we are individually (i.e. psychology) and how our individual brains work (cognitive science) with current thinking on global cultural ethics and corresponding philosophies and laws. What it means to be human can vary depending upon not just who and where we are, but also *when* we are, and how we see ourselves at any given time. When crafting ethical guidelines for AI, we must consider ethics in all forms, particularly accounting for the cultural constructs that differ between regions and groups of people in time and space.

The ethics that are informing AI and automated digital technology are essentially biased. Many of the proposals for ethics in AI that have been developed by computer scientists, engineers, politicians and other powerful entities are flawed. They neglect much of the world's many cultures and states of being and they lack the diversity in education, culture, ethnicity and gender found in today's complex world. For example, a search of the Organisation for

27 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/ethics-of-ai/256459

Related Content

Public Access ICT in South Africa

Tina James, Alan Finlay, Michael Jensen, Mark Neville and Rasagee Pillay (2012). *Libraries, Telecentres, Cybercafes and Public Access to ICT: International Comparisons* (pp. 429-451).

www.irma-international.org/chapter/public-access-ict-south-africa/55852

The Social Responsibility of Information Systems Developers

N. F. du Plooy (2003). *Socio-Technical and Human Cognition Elements of Information Systems* (pp. 41-59).

www.irma-international.org/chapter/social-responsibility-information-systems-developers/29321

Reflections on the NatureCHI Workshop Series: Unobtrusive User Experiences with Technology in Nature

Jonna Häkkinilä, Nicola J. Bidwell, Keith Cheverst, Ashley Colley, Felix Kosmalla, Simon Robinson and Johannes Schöning (2018). *International Journal of Mobile Human Computer Interaction* (pp. 1-9).

www.irma-international.org/article/reflections-on-the-naturechi-workshop-series/207699

Optimality-Theoretic Lexical Mapping Theory: A Case Study of Locative Inversion

One-Soon Her (2006). *International Journal of Technology and Human Interaction* (pp. 67-94).

www.irma-international.org/article/optimality-theoretic-lexical-mapping-theory/2879

Trusting Computers Through Trusting Humans: Software Verification in a Safety-Critical Information System

Alison Adam and Paul Spedding (2009). *Cross-Disciplinary Advances in Human Computer Interaction: User Modeling, Social Computing, and Adaptive Interfaces* (pp. 384-398).

www.irma-international.org/chapter/trusting-computers-through-trusting-humans/7298