Jared Bielby, Rachel Fischer, Geoffrey Rockwell:

Introduction to AI, Ethics & Society

The International Review of Information Ethics (IRIE) has pioneered ethics in technology and information studies for almost two decades. Leading on topics ranging from gaming to artificial intelligence, IRIE has ventured some of the most thought-provoking conversations of the digital age. As we learn to deal with Covid-19, our global information societies are set to face several new information and technology challenges, where leadership in information ethics will become more critical than ever. IRIE promises to carry forward the academic tradition of addressing the ethics of new challenges started by Dr. Rafael Capurro, founder of IRIE.

As we deal with the pandemic, the editors of IRIE are proud to announce that in the spring of 2020, IRIE successfully re-launched at the University of Alberta, Canada, where it will be hosted by the Kule Institute for Advanced Study (KIAS). The journal is now hosted on the robust open publishing infrastructure of the University of Alberta and has undergone a new design. It has also adopted a new editorial advisory board and managing editors. The re-launch of the journal accompanies the relaunch of the International Center for Information Ethics, also now supported by the University of Alberta.

The new IRIE Managing Editors team reflects a balance of expertise that both carries forward the original vision of the journal and its founders while also aligning to the academic vision of KIAS and benefiting from the University of Alberta Open Journal System. As such, IRIE is happy to introduce the following team members:

Rachel Fischer, researcher and ethicist, joins IRIE as its new Deputy Editor, and will manage the academic vision alongside IRIE Editor-in-Chief, Jared Bielby. Mihaela Ilovan, project manager and librarian, offers her technical expertise as Production Editor and will bring a professional level of quality to the production and reach of IRIE, working alongside the Open Journal System team to bring the most out of the IRIE interface. Geoffrey Rockwell, Professor of Philosophy and Digital Humanities, joins IRIE as Assistant Editor, University of Alberta Liaison, where he will manage and promote relations between the University of Alberta and the IRIE team. Lastly, but certainly not least, longstanding IRIE editors Nakada Makoto, Univ. of Tsukuba, and Johannes J Britz, University of Wisconsin-Milwaukee, join the team as Associate Editors, where they will help guide and edit forthcoming editions, offering their editor's expertise, as well as formulate new directions for future editions. Makoto and Britz have served IRIE since its inception and both bring a wealth of international and intercultural experience to the team.

IRIE looks forward to several forthcoming issues, including a special issue honouring the life and work of Norbert Wiener, an issue on Information Ethics (IE) addressing the origins and evolutions of the field, and an issue on misinformation and fake news.

The first issue on our new platform is on *Artificial Intelligence, Ethics and Society*. We bring together a set of papers that were presented at a conference with this title hosted by the Kule Institute for Advanced Study that has made AI, Ethics and Society an area of research focus. The conference was held in May of 2019. Details can be found on the KIAS web site (https://www.ualberta.ca/kule-institute/news-events/ai-ethics-and-society-conference/index.html.

Discussion of ethics in AI is set to play a critical role in the development of social license and standards for Canada as it develops an AI industry. Gartner Analytics top ten strategic technology trends for 2019 include both AI-driven development and Digital Ethics among the ten leading trends of 2019, highlighting the growing concern and focus on ethics for AI over the last number of years (Gartner).

In June of 2018, Canadian Prime Minister Justin Trudeau and French President Emmanuel Macron announced the signing of the Canada-France Statement on Artificial Intelligence (Canada-France Statement). This statement reaffirmed the G7 Innovation Ministers' Statement on Artificial Intelligence from March 2018 (G7 Innovation). In the Canada-France Statement both countries committed to promoting "a vision of human-centric artificial intelligence grounded in human rights, inclusion, diversity, innovation and economic growth."

Trudeau's commitment to human-centric AI was part of a larger federal initiative supporting investment in AI research through the Canadian Institute for Advanced Research (CIFAR) via a Pan-Canadian Artificial Intelligence Strategy. The strategy supports three new AI institutes across Canada including Amii (Alberta Machine Intelligence Institute) in Edmonton. The University of Alberta is thus one of the three lead AI research hubs in Canada supported by the Pan-Canadian Artificial Intelligence Strategy.

The responsible development and deployment of AI is an important issue for Canadians. When the Prime Minister made further funding announcements at the G7 Multistakeholder Conference on AI in Montréal in December of 2018 they followed an agenda on Enabling the Responsible Adoption of AI (Government of Canada). Further, there have been two important meetings in Canada that have led to declarations around responsible AI. The Montreal Declaration for a Responsible Development of Artificial Intelligence was developed through a participatory process that included a forum in November 2017 (Montréal Declaration). The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning system was launched at a forum organized by Access Now in May of 2018 (Access Now).

In light of these commitments, the Kule Institute for Advanced Study (KIAS) at the University of Alberta organized an interdisciplinary and intersectoral conference in May of 2019. The conference brought together researchers in the history, ethics, policy, business and science of AI with stakeholders from business and government. The goal of the conference was to mobilize knowledge about the social and ethical impact of AI in a fashion that encourages innovation that is open, safe and beneficial to all Canadians. The Edmonton conference on *AI, Ethics and Society* brought humanists and social scientists into the discussion that was until recently dominated by technologists alone. The conference, through partnerships with the Alberta Machine Intelligence Institute (Amii), ATB Financial, and Compute Canada connected AI researchers in computing science with social science and humanities researchers and stakeholders from business and government. Conference participants were drawn from both local and international sources, including tapping into the vast resources and talent available at the University of Alberta so as to build a community of experts in AI ethics to mirror and partner with the scientific community. International partners like the IEEE (Institute of Electrical and Electronic Engineers) and ICIE (International Center for Information Ethics) brought an international track record of developing insights and standards for industry, academia and society.

The University of Alberta is a global leader in AI research and development; as partner of the ICIE and host of IRIE it is both timely and advantageous for the first issue of the relaunched IRIE journal to focus on AI Ethics. While Canadian universities like the University of Alberta lead in AI research, researchers have yet to develop a parallel dialogue around the ethical and social impacts of big data and AI technologies that can engage the science and technology rather than follow it. The *International Review of Information Ethics* has chosen to publish selected conference papers in the following Special Issue. Our hope is that this Special Issue on *AI, Ethics and Society* will further engage the social sciences and humanities to engage their colleagues working in AI in a dialogue.

Rafael Capurro's paper considers issues that have arisen when dealing with societal and ethical implications of AI since the seventies. He reflects whether machines can think, what AI as distributed intelligence means and the connectedness between artificial things and living agents. He concludes by presenting some challenges and findings resulting from academia, scientific and political institutions.

Trust, privacy and transparency are some of the information ethical challenges that arise in this edition. Hongladarom and Jinnie Shin, Okan Bulut, Mark J. Gierl address trust in their two articles. The former prioritises trust and reputation, arguing that safeguards are needed to guarantee privacy rights. Hongladarom claims that these values and related considerations must be taken seriously due to the globalized nature of both ICTs and ethics. The article by Shin, Bulut and Gierl addresses the concern by focussing on Canadian data scientists' self-confidence in creating trusted AI systems. Hence, although AI has benefits for society, it is still questionable whether a trusting relationship can exist between AI and humans.

Some of these benefits relate to improved service delivery and practices of innovation within higher education. Ipperciel explores the notion of student centeredness and innovation by looking at a project in which an AI-

powered Student Virtual Assistant was developed at York University in Toronto. This is complemented by Potgieter's article on "Privacy concerns in educational data mining and learning analytics". He investigates these techniques which aim to utilise data processing to inform all aspects of educational practice.

In Canada, privacy regulators have seen massive information sharing projects fail when public input or consultation, or independent oversight is not considered. Clayton and Sibbald provide an overview of changes that have occurred in practice and how these regulators have started to incorporate discussions on ethics.

Ingram endeavours to shed light on the black box of AI. Transparent practices need to be introduced in automated processes such as being considered for a job or bank loan. She examines a body of literature related to the ethical issues of fairness, accuracy and transparency. Cohn further elaborates on this when arguing that much of the rationale for developing these technologies focus on increasing scalability/optimization and innovation. What will our actions and discourses look like when we place care at the centre of our ethical reflections? To achieve such thinking, Berendt proposes the concept of Ethics Pen-Testing (EPT). The EPT identifies challenges and pitfalls when determining, from an AI point of view, what the Common Good is and how it can be enhanced by AI.

Finally, Nye philosophizes on AI and workers' living standards where he argues the value in left-wing and right-wing political philosophy, engaging John Rawls's Justice as Fairness and Robert Nozick's Entitlement Theory to demonstrate how political action is needed to counteract the effects of technological displacement.

This edition considers themes from ethical best practices for industry and government developing responsible AI services to aligning cultural and societal values in AI design, the role of researchers from social sciences and humanities disciplines in ethical innovation in the AI sector, and methods for interdisciplinary and intersectoral collaboration between interested in responsible AI. It looks at how nations can contribute to building a healthy AI sector through policy, research and innovation.

References

Access Now (2018). "The Toronto Declaration: Protecting the rights to equality and nondiscrimination in machine learning systems." <https://www.accessnow.org/the-torontodeclaration-protecting-the-rights-to-equality-and-non-discrimination-in-machine-learningsystems/>

Canada-France Statement on Artificial Intelligence. <https://international.gc.ca/world-monde/international_relations-relations_internationales/europe/2018-06-07-france_ai-ia_france.aspx?lang=eng>

G7 Innovation Ministers' Statement on Artificial Intelligence. <http://www.g8.utoronto.ca/employment/2018-labour-annex-b-en.html>

Gartner Identifies the Top 10 Strategic Technology Trends for 2019. Retrieved June 29, 2020. https://www.gartner.com/en/newsroom/press-releases/2018-10-15-gartner-identifies-the-top-10-strategic-technology-trends-for-2019

Government of Canada (2018). "G7 Multistakeholder Conference on Artificial Intelligence." <http://www.ic.gc.ca/eic/site/133.nsf/eng/home>

"Montréal Declaration: Responsible AI." <https://www.montrealdeclarationresponsibleai.com/the-declaration>