WASP-AI course - Artificial Intelligence and Society: on the legal, ethical and societal aspects of AI

Before the start of the course students are required to read all the literature marked with (*) and strongly advised to read the other articles indicated for each lecture.

Lectures

1. Responsible Artificial Intelligence – Professor Virginia Dignum (Umeå University)

Every day we see news about advances and the societal impact of AI. AI is changing the way we work, live and solve challenges but concerns about fairness, transparency or privacy are also growing. What is responsible AI and how can we ensure that AI is used and developed in a responsible way? Ensuring an ethically aligned purpose is more than designing systems whose result can be trusted. It is about the way we design them, why we design them, and who is involved in designing them. If we are to produce responsible trustworthy AI, we need to work towards technical and socio-legal initiatives and solutions which provide concretise instructions, tools, and other means of dictating, helping, and educating AI practitioners at aligning their systems with our societies' principles and values.

Readings:

(*) chapter 4 of Responsible Artificial Intelligence: how to develop and use AI in a responsible way (attached).

About the lecturer:

Virginia Dignum is Professor of Social and Ethical Artificial Intelligence at Umeå University, Sweden and associated with the TU Delft in the Netherlands. She is the scientific director of WASP-HS, the Wallenberg Program on Humanities and Society for AI, Autonomous Systems and Software. She is a Fellow of the European Artificial Intelligence Association (EURAI), a member of the European Commission High Level Expert Group on Artificial Intelligence, of the World Economic Forum's Global Artificial Intelligence Council, of the Executive Committee of the IEEE Initiative on Ethically Aligned Design, and a founding member of ALLAI-NL, the Dutch AI Alliance. She is the author of "Responsible Artificial Intelligence: developing and using AI in a responsible way" published by Springer in 2019. She has a PHD in Artificial Intelligence from Utrecht University and in 2006 she was awarded the prestigious Veni grant by the NWO (Dutch Organization for Scientific Research).

2. Introduction to Ethics of AI – Dr. Aimee van Wynsberghe (Delft University of Technology)

Introduction to Ethics of AI is designed to introduce non-philosophers to the various themes found in the ethics of AI, drawing on the ethics of technology. The session will provide an overview of the study of ethical theories and some of the key ethical issues related to AI currently discussed in academia. In this course the following kinds of questions are examined: What are the main ethical concepts guiding the ethics of AI? How can one address the ethical issues raised through AI? What risks and uncertainties need to be accounted for when deploying AI? Who is responsible for mitigating and/or preventing societal risks associated with AI?

Readings:

- (*) Poel, I. van de. (2013). Why New Technologies Should be Conceived as Social Experiments. *Ethics, Policy & Environment*, *16*(3), 352–355. (link)
- (*) van Wynsberghe, Aimee, & Robbins, S. (2014). Ethicist as designer: A pragmatic approach to ethics in the lab. *Science and Engineering Ethics*, *20*(4), 947–961.   (link)
- Johnson, D. G. (2015). Technology with No Human Responsibility? *Journal of Business Ethics*, *127*(4), 707–715. (link)
- van Wynsberghe, A. (2012). Designing Robots for Care: Care Centered Value-Sensitive Design. *Science and Engineering Ethics*, *19*(2), 407–433. (link)
- Floridi, L. (2019). Translating principles into practices of digital ethics: five risks of being unethical. *Philosophy & Technology*, *32*(2), 185-193. (attached)

About the lecturer:

Aimee van Wynsberghe has been working in ICT and robotics since 2004. She began her career as part of a research team working with surgical robots in Canada at CSTAR (Canadian Surgical Technologies and Advance Robotics). She is Assistant Professor in Ethics and Technology at TU Delft in the Netherlands. She is co-founder and co-director of the Foundation for Responsible Robotics and on the board of the Institute for Accountability in a Digital Age. She is a 2018 L'Oreal Unesco 'For Women in Science' laureate. Aimee also serves as a member of the European Commission's High-Level Expert Group on AI and is a founding board member of the Netherlands AI Alliance. Aimee has been named one of the Netherlands top 400 influential women under 38 by VIVA and was named one of the 25 'women in robotics you need to know about'. She is author of the book *Healthcare Robots: Ethics, Design, and Implementation* and has been awarded an NWO personal research grant to study how we can responsibly design service robots. She has been interviewed by BBC, Quartz, Financial Times, and other International news media on the topic of ethics and robots, and is often invited to speak at International conferences and summits.

3. Legal Issues in AI - Professor Mireille Hildebrandt (VUB, Brussels)

In this lecture we will probe the difference between legal and other 'by design' approaches, starting from the fact that nobody is above the rule of law - not even developers of computational systems. This means that before defining your own concepts of privacy, fairness, bias or freedom of expression it is important to understand the corresponding fundamental rights as they must be respected at the level of the design, at least in the case that EU data protection law applies. We will do some close reading into the legal requirements of Data Protection by Design and Default (LPbDD) and Data Protection Impact Assessment (DPIA), taking note that to grasp the meaning of these two 'by design' approaches we have to situate them in the framework of the rule of law. The rule of law does not refer to a bag of rules but to a coherent, dynamic system of checks and balances that ensures countervailing powers, which should also be operational at the level of research design, technical infrastructures and applications.

Readings:

(*) chapters 5 and 10 of Law for Computer Scientists and Other Folk, which is coming out beginning April, but available in open access at: https://lawforcomputerscientists.pubpub.org.

About the Lecturer:

Mireille Hildebrandt is a Research Professor on 'Interfacing Law and Technology' at Vrije Universiteit Brussels (VUB), appointed by the VUB Research Council. She is co-Director of the Research Group on [Law Science Technology and Society studies (LSTS)](#) at the Faculty of Law and Criminology. She also holds the part-time Chair of Smart Environments, Data Protection and the Rule of Law at the Science Faculty, at the [Institute for Computing and Information Sciences (iCIS)](#) at Radboud University Nijmegen. Her research interests concern the implications of automated decisions, machine learning and mindless artificial agency for law and the rule of law in constitutional democracies. Hildebrandt has published 4 scientific monographs, 23 edited volumes or special issues, and over 100 chapters and articles in scientific journals and volumes. She received an ERC Advanced Grant for her project on 'Counting as a Human Being in the era of Computational Law' (2019-2024).

4. Social Artificial Intelligence – Professor Frank Dignum (Umeå University)

At the moment, many initiatives for sustainable cities concentrate on the ecological aspect and try to optimize on one aspect, such as transport, energy or waste management. However, all these aspects are related and an optimal solution found for one aspect might be bad for another aspect. E.g. stimulating electric cars by creating free charging stations at work can increase the number of electric cars, but prevent people going to work by bike.

One of the pittfalls of current approaches to optimizing environmental sustainable behavior is that they assume the preferences of people to be stable. But, people are both social and habitual animals. I.e. they will change their preference based on what they sense other people's preferences are. If many people start going to work on the bike and it becomes a status symbol for young urban people then suddenly people might change their preference in order to also belong to this desirable group. Thus when looking for solutions to optimize the sustainability of a city one should keep in mind that people will adapt to the changing situation and can change preferences based on experiences and circumstances.

In my presentation I will argue that we should be using AI in its social context:

1. Provide an **integral environment** in which cities aiming for a sustainable future can experiment and simulate integrated solutions for different aspects of sustainable behavior.
2. Provide **fundamental theoretical** approaches that can be used to determine optimal solutions in a dynamic and adaptive society
3. Provide **practical support** for cities that want to take the first steps on the road to smart sustainability based on their current situation, context and budget requirements.

Readings:

- (*) Dignum, Frank, et al. "A conceptual architecture for social deliberation in multi-agent organizations." *Multiagent and Grid Systems* 11.3 (2015): 147-166. ([link](#))
- (*) Heidari, Samaneh, Maarten Jensen, and Frank Dignum. "Simulation with Values." *Social Simulation Conference.* 2018. ([link](#))
- Nooitgedagt, Aldert, Robbert Jan Beun, and Frank Dignum. "e-Coaching for intensive cardiac rehabilitation." *International Conference on Persuasive Technology.* Springer, 2017 ([link](#))

About the lecturer:

Frank Dignum received his PhD at the VU in Amsterdam in 1989. Since 2019 he is Wallenberg distinguished chair in AI at Umeå University in Sweden. He also has an affiliation to Utrecht University still and since 2013 he is honorary principal research fellow of the University of

Melbourne. Since 2014 he is a EurAI fellow. He is well known for his work on norms and his theory of social agents is employed in social simulations to support policy making and e-coaching. He published 18 books and more than 300 papers. At this moment his H-index is 54.