

AI FROM PHILOSOPHY TO SCIENCE TO TECHNOLOGY TO ETHICS TO LAW AND BACK

PART 4

AI ETHICS INITIATIVES & LAWS

GORDANA DODIG-CRNKOVIĆ
CHALMERS UNIVERSITY OF TECHNOLOGY
GOTHENBURG, SWEDEN

Ethical Considerations in AI
Gothenburg, 2020

<http://gordana.se/Presentations>



<https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/Ai-genie-in-a-bottle>

CONTENT

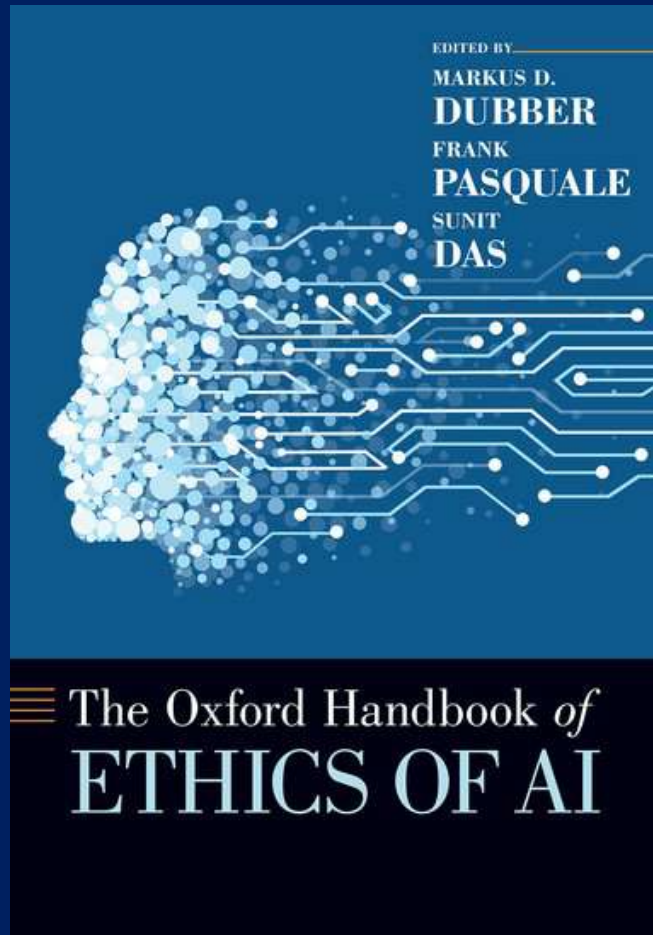
Previous parts of the lecture covered the following:

Part 1 philosophical and transdisciplinary scientific view of AI

Part 2 technological overview and value-based foundations of AI

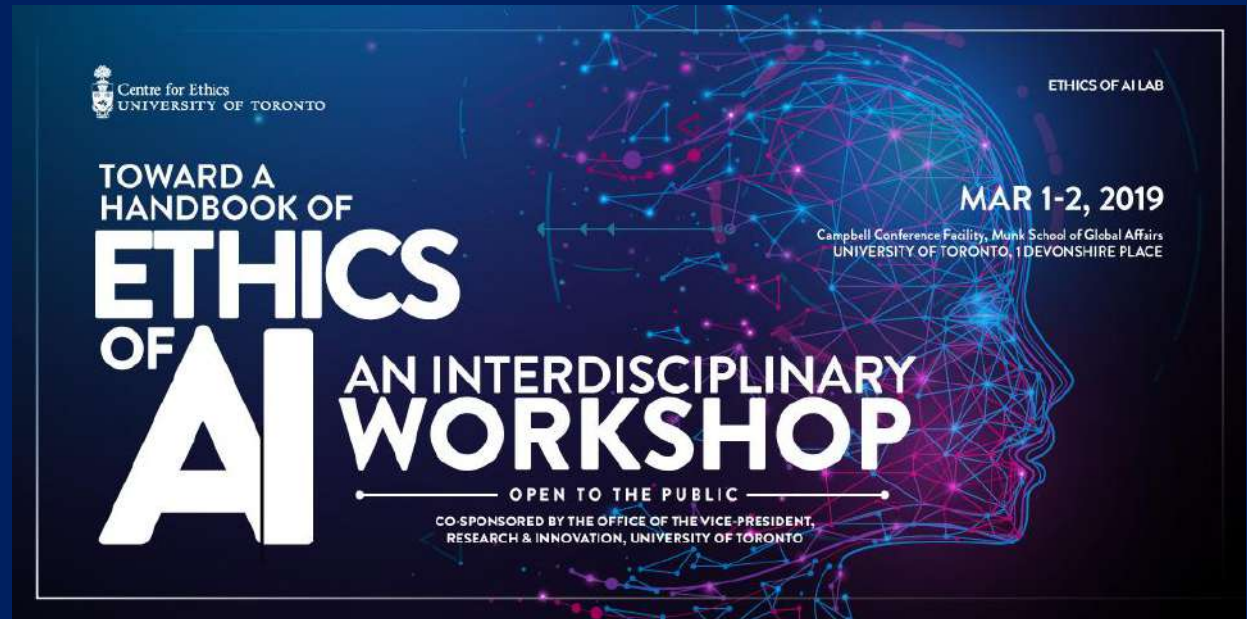
Part 3 ethics of AI represented by guidelines and policies

Now we proceed to the topic of the state of the art of the ethics of AI as represented by various initiatives and laws. We finish with conclusions from the Part1-Part4.



ONGOING INITIATIVES

The Oxford Handbook of Ethics of AI



<https://global.oup.com/academic/product/the-oxford-handbook-of-ethics-of-ai-9780190067397?cc=ca&lang=en&#>

The Oxford Handbook of Ethics of AI

Part I. Introduction & Overview

1. The Artificial Intelligence of Ethics of AI: An Introductory Overview
2. The Ethics of Ethics of AI: Mapping the Field
3. Ethics of AI in Context: Society & Culture

Part II. Frameworks & Modes

4. Why Industry Self-regulation Will Not Deliver 'Ethical AI': A Call for Legally Mandated Techniques of 'Human Rights by Design'
5. Private Sector AI: Ethics and Incentives
6. Normative Modes: Codes & Standards
7. Normative Modes: Professional Ethics

Part III. Concepts & Issues

8. Fairness and the Concept of 'Bias'
9. Accountability in Computer Systems
10. Transparency
11. Responsibility
12. The Concept of Handoff as a Model for Ethical Analysis and Design
13. Race and Gender
14. The Future of Work in the Age of AI: Displacement, Augmentation, or Control?
15. The Rights of Artificial Intelligences
16. The Singularity: Sobering up About Merging with AI
17. Do Sentient AIs Have Rights? If So, What Kind?
18. Autonomy
19. Troubleshooting AI and Consent
20. Judgment, Error, and Authority in the Codification of Law

IV. Perspectives & Approaches

22. Computer Science
23. Engineering
24. Designing Robots Ethically Without Designing Ethical Robots: A Perspective from Cognitive Science
25. Economics
26. Statistics
27. Automating Origination: Perspectives from the Humanities
28. Philosophy
29. The Complexity of Otherness: Anthropological contributions to robots and AI
30. Calculative Composition: The Ethics of Automating Design
31. Global South
32. East Asia
33. Artificial Intelligence and Inequality in the Middle East: The Political Economy of Inclusion
34. Europe's struggle to set global AI standards

Part V. Cases & Applications

35. The Ethics of Artificial Intelligence in Transportation
36. Military
37. The Ethics of AI in Biomedical Research, Medicine and Public Health
38. Law: Basic Questions
39. Law: Criminal Law
40. Law: Public Law & Policy: Notice, Predictability, and Due Process
41. Law: Immigration & Refugee Law
42. Education
43. Algorithms and the Social Organization of Work
44. Smart City Ethics

HUMAN-CENTERED ARTIFICIAL INTELLIGENCE

We are designing the principles for a new science that will make artificial intelligence based on European values and closer to Europeans.

This new approach works toward AI systems that augment and empower all Humans by understanding us, our society and the world around us.

Result: Research Roadmap >

Result: Policy Guidelines >

Result: Ethics Framework >

Result: Connecting Communities >

Result: Dynamic Funding >

Result: Micro Projects >

FEATURED EVENT

HumaneAI delivering a one day event @ European Parliament

Presenting the new science of Artificial Intelligence with European values

Learn more >

HUMAN-CENTRIC HUMANE AI

"The goal of Humane AI is to harness the emergence of enabling technologies for human-level interaction to empower individuals and society, by providing new abilities to perceive and understand complex phenomena, to individually and collectively solve problems, and to empower individuals with new abilities for creativity and experience."

<https://www.humane-ai.eu>



European Commission > Strategy > Digital Single Market > Policies >

Digital Single Market

POLICY

High-Level Expert Group on Artificial Intelligence

<https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence>

52 experts of a High-Level Expert Group on Artificial Intelligence, are comprising representatives from academia, civil society, as well as industry, supporting the implementation of the European Strategy on Artificial Intelligence.

Ethics Guidelines on Artificial Intelligence: The Guidelines put forward a human-centric approach on AI and list 7 key requirements that AI systems should meet in order to be trustworthy.

1. Human agency and oversight
2. Technical robustness and safety
3. Privacy and Data governance
4. Transparency
5. Diversity, non-discrimination and fairness
6. Societal and environmental well-being
7. Accountability

European Strategy on AI



The screenshot shows a web browser window with the URL ec.europa.eu/digital-single-market/en/artificial-intelligence. The page header includes the European Commission logo and a breadcrumb trail: "European Commission > Strategy > Shaping Europe's digital future > Policies >". Below this, the text "Shaping Europe's digital future" and "POLICY" are visible. The main heading is "Artificial Intelligence". Underneath, there is a "PAGE CONTENTS" section with the following items: "A European approach to Artificial Intelligence", "Coordinated Plan on Artificial Intelligence 'Made in Europe'", "Building Trust in Human-Centric Artificial Intelligence", "Declaration of cooperation on Artificial Intelligence", and "Useful links".

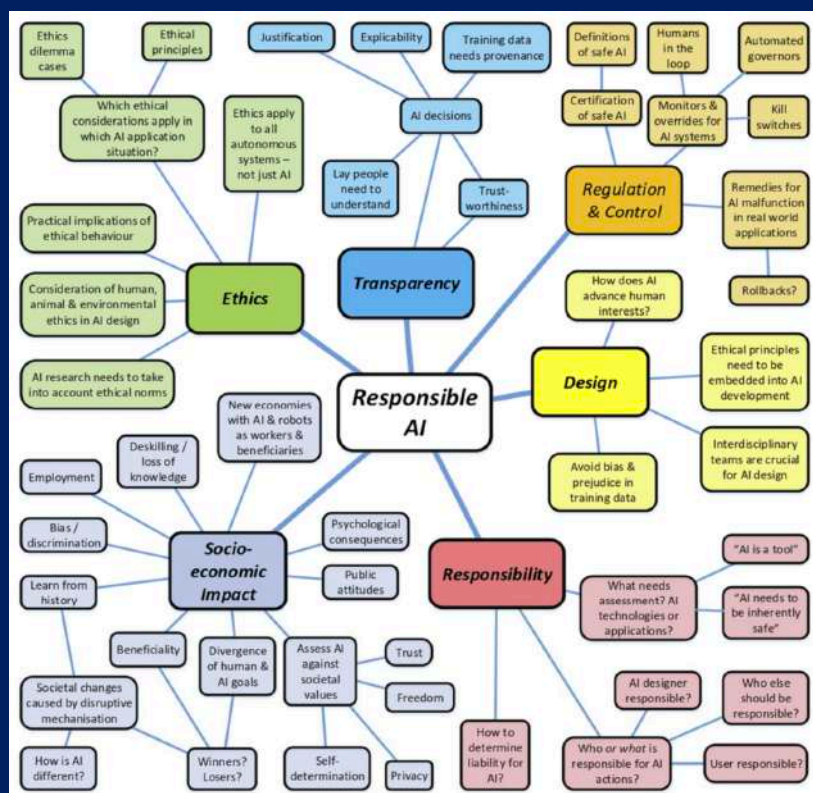
<https://ec.europa.eu/digital-single-market/en/artificial-intelligence>

RESPONSIBLE AI (NOT CAUSING HARM) & AI FOR GOOD (ACTIVELY CONTRIBUTING TO HUMANITY)

“Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation” Report by cross-disciplinary experts.

“Responsible AI” is an umbrella term for investigations into legal, ethical and moral standpoints of autonomous algorithms or applications of AI whose actions may be safety-critical or impact the lives of citizens in significant and disruptive ways. It reports a summary of results from a consultation with cross-disciplinary experts in and around the subject .”

Steve Taylor, Brian Pickering, Michael Boniface, University of Southampton, UK
 Michael Anderson, Uni. of Hartford, USA
 David Danks, L.L., Carnegie Mellon USA
 Dr Asbjørn Følstad, SINTEF, Norway
 Dr. Matthias Leese, ETH Zurich, CH
 Vincent C. Müller, University of Leeds, UK
 Tom Sorell, University of Warwick, UK
 Alan Winfield, Uni. of the West of England
 Dr Fiona Woollard, Uni. Southampton, UK



<https://www.ngi.eu/news/2018/07/23/responsible-ai/>



<https://ai4good.org/>
https://en.wikipedia.org/wiki/AI_for_Good
<https://www.microsoft.com/en-us/ai/ai-for-good>

WHITE PAPER ON AI ECOSYSTEM

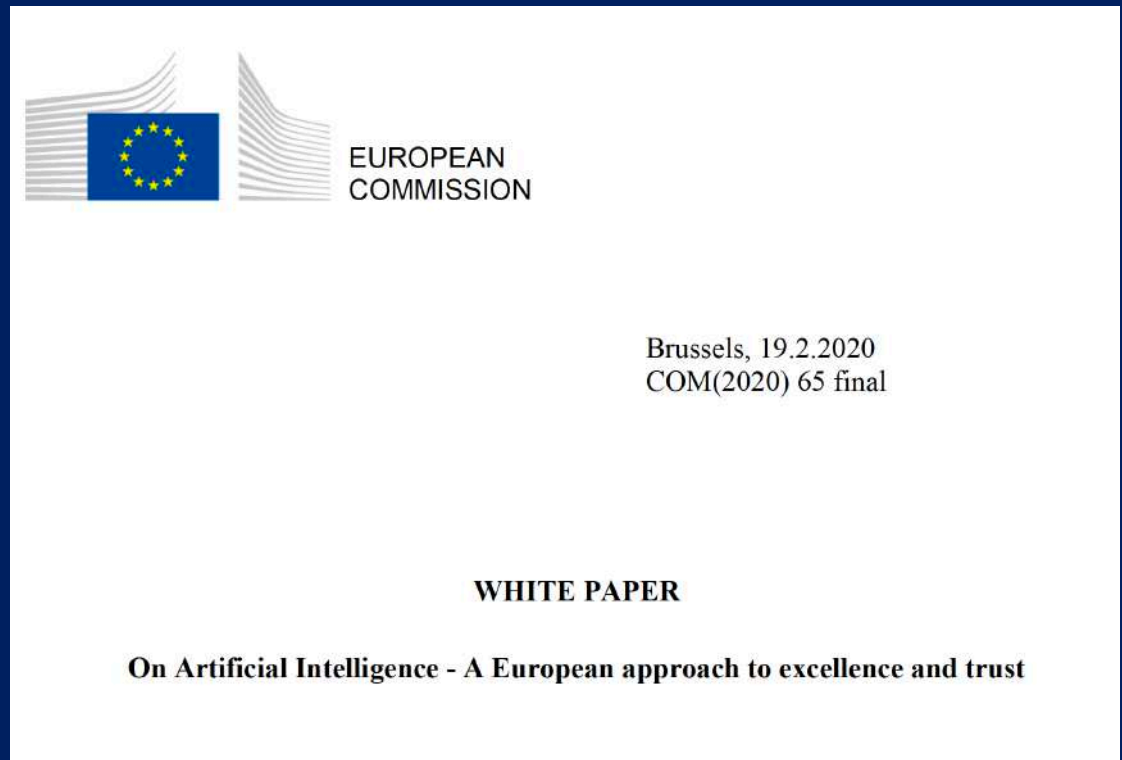
Comments invited until 19 May 2020

- AI definition
- AI context (socio-technological system)
- “AI Race” versus “AI Exploration”
- Trustworthy AI (cannot be a choice between an accurate black box AI-system or an explainable, but less accurate AI-system)
- Explainability
- Bias and transparency
- Traceability on the decisions made by the human actors related to the design, development, and deployment of a system
- Liability (requires adjustments to the existing safety and liability regimes)
- Assessment for high-risk AI based on Ethics Guidelines for Trustworthy AI, developed by the High-Level Expert Group on AI.

Comments invited until 19 May 2020

<http://allai.nl/first-analysis-of-the-eu-whitepaper-on-ai/>

First analysis of the EU Whitepaper on AI



https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

LEGAL VIEW OF AI

- AI develops faster than laws
- Liability in case AI causes damage or loss of lives, like autonomous weapons
- When artificial intelligence is not enough but common sense is needed. Biases and errors.
- Privacy loss because of data-hungry AI
- Patents based on or produced by AI, intellectual Property
- Job loss and wealth inequality
- Robot rights. How should we treat AIs (when they get more intelligent)



LEGAL ASPECTS OF AI

Mireille
Hildebrandt



- Law as architecture
- The choice architecture of the Rule of law
- The GDPR and the Charter of Fundamental Rights
- The methodological integrity of computer science and the GDPR
- Legal protection by design

29/10/19

ECSS 2019 ROME Keynote Hildebrandt

Mireille Hildebrandt is a Dutch lawyer and philosopher who works at the intersection between law and computer science, the Research Professor on 'Interfacing Law and Technology' at the Vrije Universiteit Brussel. Principal investigator of the 'Counting as a Human Being in the Era of Computational Law' project. [Wiki]

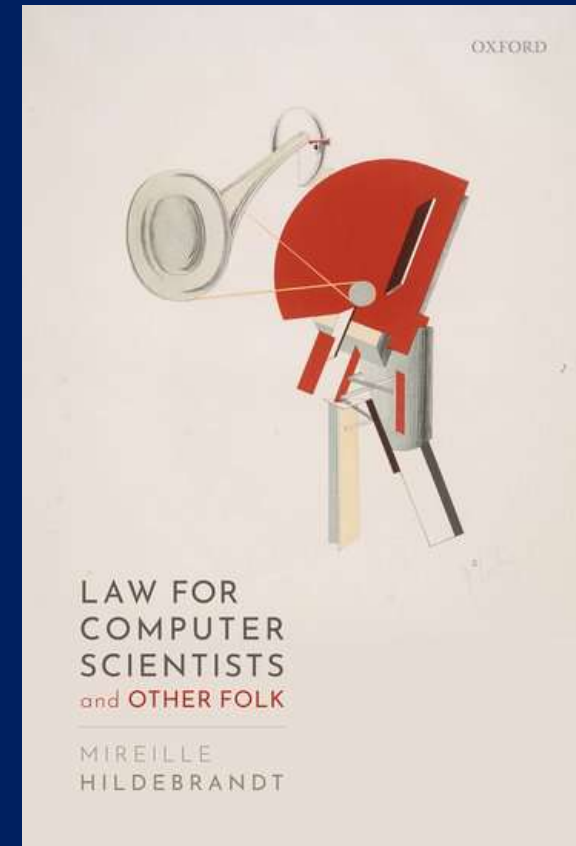
LAW FOR COMPUTER SCIENTISTS

Mireille Hilderbrandt

Forthcoming 2020 by Oxford University Press
Available online, also for comments
E-book open access

<https://lawforcomputerscientists.pubpub.org/>

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme on 'Counting as a Human Being in the Era of Computational Law'



LAW FOR Computer SCIENTISTS

Mireille Hilderbrandt

Part I

What Law Does

1. Introduction: Textbook and Essay
2. Law, Democracy, and the Rule of Law
3. Domains of Law: Private, Public, and Criminal Law
4. International and Supranational Law

Part II

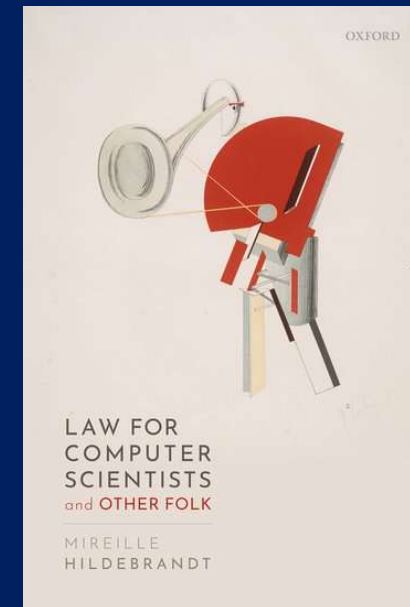
Domains of Cyberlaw

5. Privacy and Data Protection
6. Cybercrime
7. Copyright in Cyberspace
8. Private Law Liability for Faulty ICT

Part III

Frontiers of Law in an Onlife World

9. Legal Personhood for AI?
10. 'Legal by Design' or 'Legal Protection by Design'?
11. Closure: on ethics, code and law



<https://lawforcomputerscientists.pubpub.org/>

GDPR - General Data Protection Regulation

for citizens of the European Union

1. **Data consent:** A company that collects data on individuals must have "unambiguous" consent from those individuals — silence, pre-ticked boxes, or inactivity do not count as consent.
2. **Data portability:** Companies must be willing to move personal data to another location or company, even a direct competitor, if requested by the consumer.
3. **Data deletion:** Companies must delete personal data when requested by an individual.
4. **Consumer profiling:** Individuals can contest, object to, and request explanation for automated decisions or decisions made by algorithms.
5. **Data protection:** The GDPR has strict, specific data security requirements, and stronger enforcement. Data encryption is especially important.
6. **Data breach notification:** The GDPR has a specific definition for what constitutes a breach of "personal" data, along with strict requirements for notifying affected individuals if a breach occurs.
7. **Data Protection Officer (DPO):** All companies that store or process large amounts of personal data must appoint or hire a data protection officer (DPO), who will drive data security and oversee GDPR compliance.

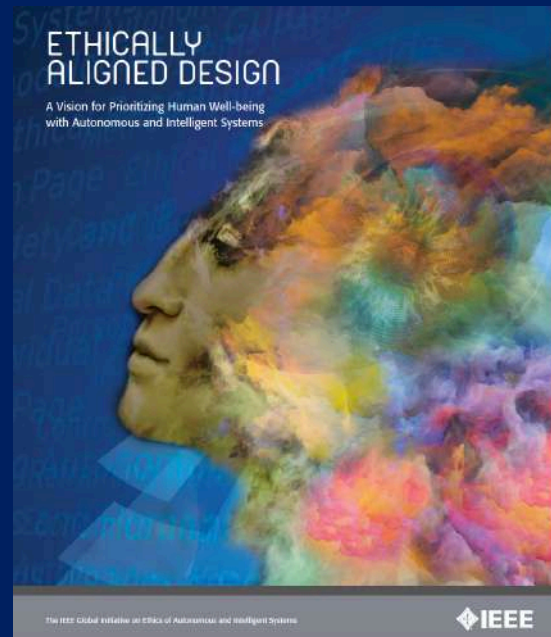
There are two tiers of fines under the GDPR.

First tier: 2% of a company's annual revenue or €10 million, whichever is **larger**.

Second tier: 4% of a company's annual revenue or €20 million, whichever is **larger**.

<https://gdpr-info.eu/>

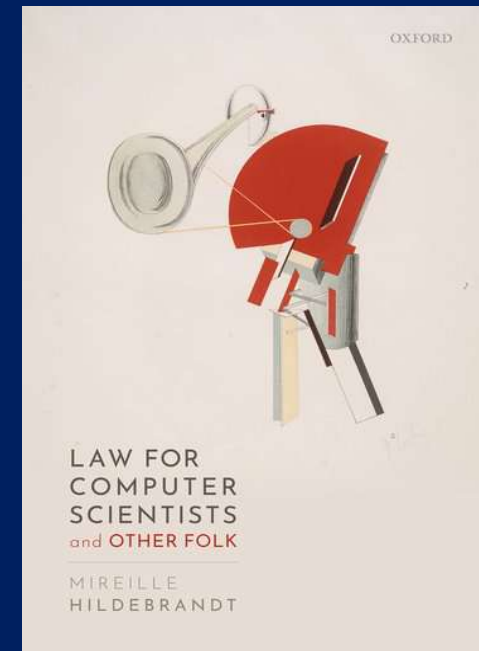
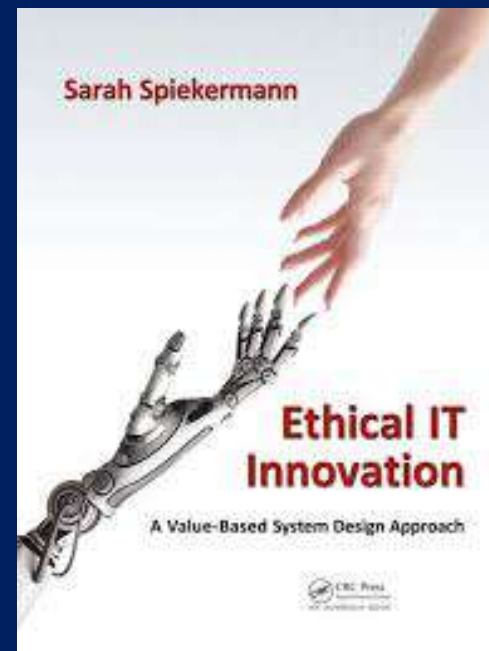
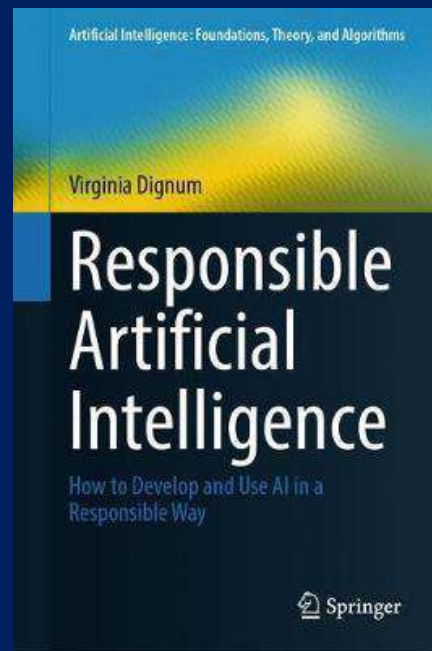
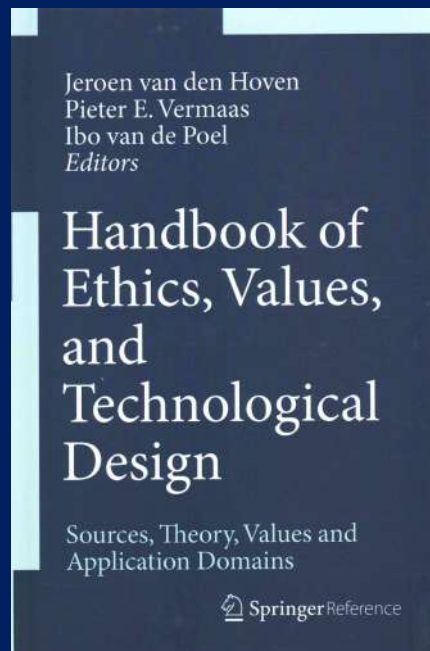
GUIDELINES, RECOMMENDATIONS, POLICY



<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

<https://ethicsinaction.ieee.org/>

RECENT BOOKS



CONCLUSION

Currently AI is in the flow - very rapid development

Learning AI

From Philosophy to Science to Technology to Law to Ethics and Back

AI can be understood from philosophical foundations - with **ontological** question about **what AI is today** and what it could be (and should be) **developed into**, its **epistemological** basis – how does AI affect our possibility to know, and **axiological/ethical** analysis of **what is good AI**.

The next step is the development of various scientific domains relevant for AI, followed tightly of its technological applications..

Legal regulations follow as soon as reasonably possible when technological applications become widespread and need legal regulation.

After practical experiences with current technology, the next step in the development is made and the **cycle (spiral) starts again**.



REFERENCES

1. European Commission's High-Level Expert Group on Artificial Intelligence. Draft Ethics Guidelines for Trustworthy AI (2019) Available online: <https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai>
2. Jobin, A., Ienca, M. and Vayena, E. (2019) The global landscape of AI ethics guidelines. Nature Machine Intelligence | VOL 1 | SEPTEMBER 2019 | 389–399 | www.nature.com/natmachintell <https://www.nature.com/articles/s42256-019-0088-2.pdf>
3. WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust. Brussels, 19.2.2020. COM(2020) 65 final https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
4. Spiekermann S. (2015) Ethical IT Innovation: A Value-Based System Design Approach. Taylor & Francis
5. Virginia Dignum (2019) Responsible Artificial Intelligence. How to Develop and Use AI in a Responsible Way. Springer Nature Switzerland AG
6. Asilomar Conference 2017. Asilomar AI Principles. Available online: <https://futureoflife.org/ai-principles/?cn-reloaded=1>
7. European Group on Ethics in Science and New Technologies (2018) Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. Available online: https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

REFERENCES

6. Floridi, L., Cowls, J., King, T.C. et al. How to Design AI for Social Good: Seven Essential Factors. *Sci Eng Ethics* (2020). <https://doi.org/10.1007/s11948-020-00213-5>
7. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F.; et al. (2018) AI4People—An Ethical Framework for a Good AI Society. *Minds Mach.* 28, 689–707. <https://link.springer.com/article/10.1007%2Fs11023-018-9482-5>
8. Floridi, L. (2019) Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology*. <https://doi.org/10.1007/s13347-019-00354-x>
9. Morley, J., Floridi, L., Kinsey, L., Elhalal, A. (2019) From What to How: An Overview of AI Ethics Tools, Methods and Research to Translate Principles into Practices. arXiv:1905.06876
10. Wachter S, Mittelstadt B, Floridi L (2017) Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law*, vol. 7, issue 2 (2017) pp. 76-99 Published by Oxford University Press (OUP)
11. Dodig-Crnkovic G. and Çürüklü B., Robots - Ethical by Design, *Ethics and Information Technology* 2011, Volume 14, Number 1, pp. 61-71. <http://www.springerlink.com/content/f432g33181787u63/fulltext.html>