

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

PART 3

AI ETHICS – GUIDELINES AND POLICIES

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

Part 2 technological overview and value-based foundations of AI

Now we proceed to the ethics of AI represented internationally by guidelines and policies

ETHICAL GOVERNANCE OF AI

Requirement of transparency and democratic legitimacy

Following responsible research & innovation (Europe)*

IEEE initiative – starting with education, ethics in the whole process of engineering

New types of problems emerge in “automated public sphere”

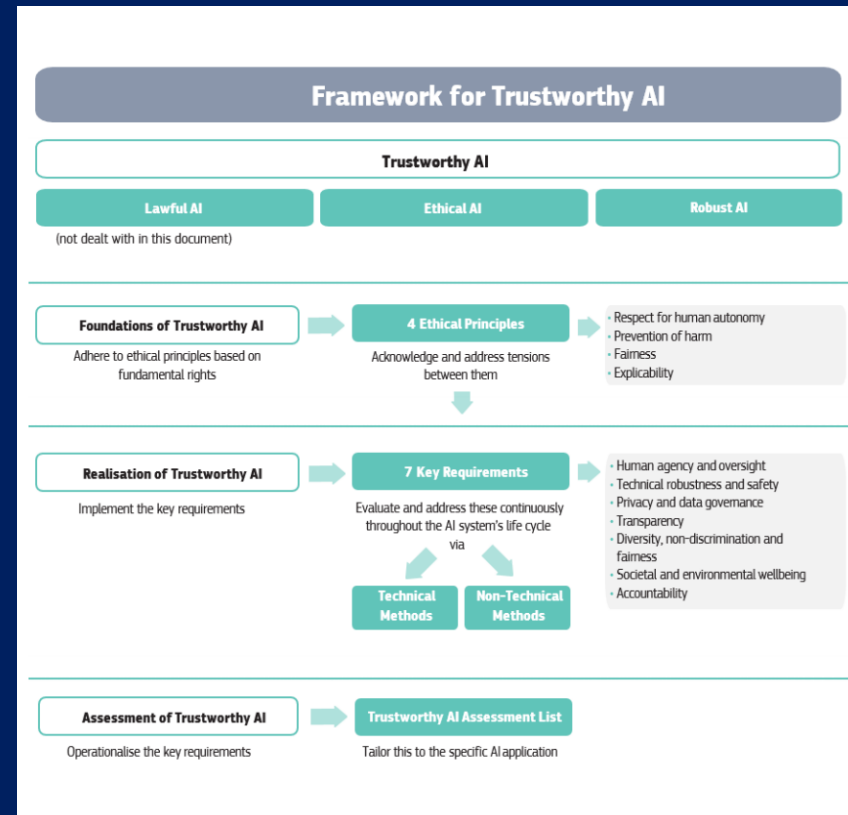
Stakeholders involvement is essential

Currently: governmental bodies are working on both legislative and guidelines

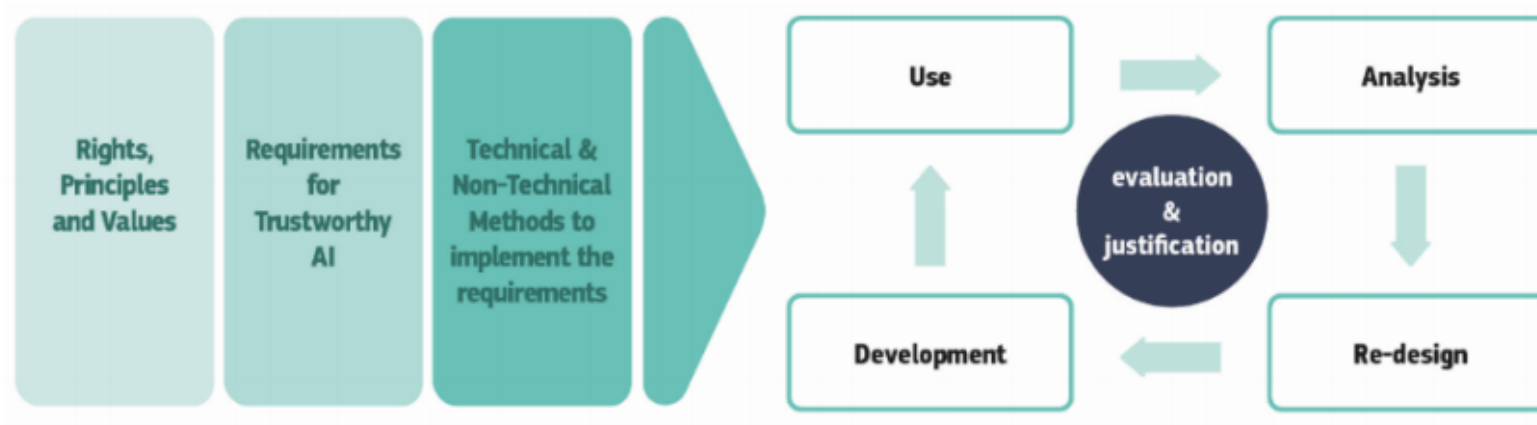
- Various groups and initiatives are proposing policies and guidelines
- Concentration of power in few companies (Google, Apple, Amazon, Microsoft) who provide their proposals for ethical policies, guidelines and practices

*<https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation>

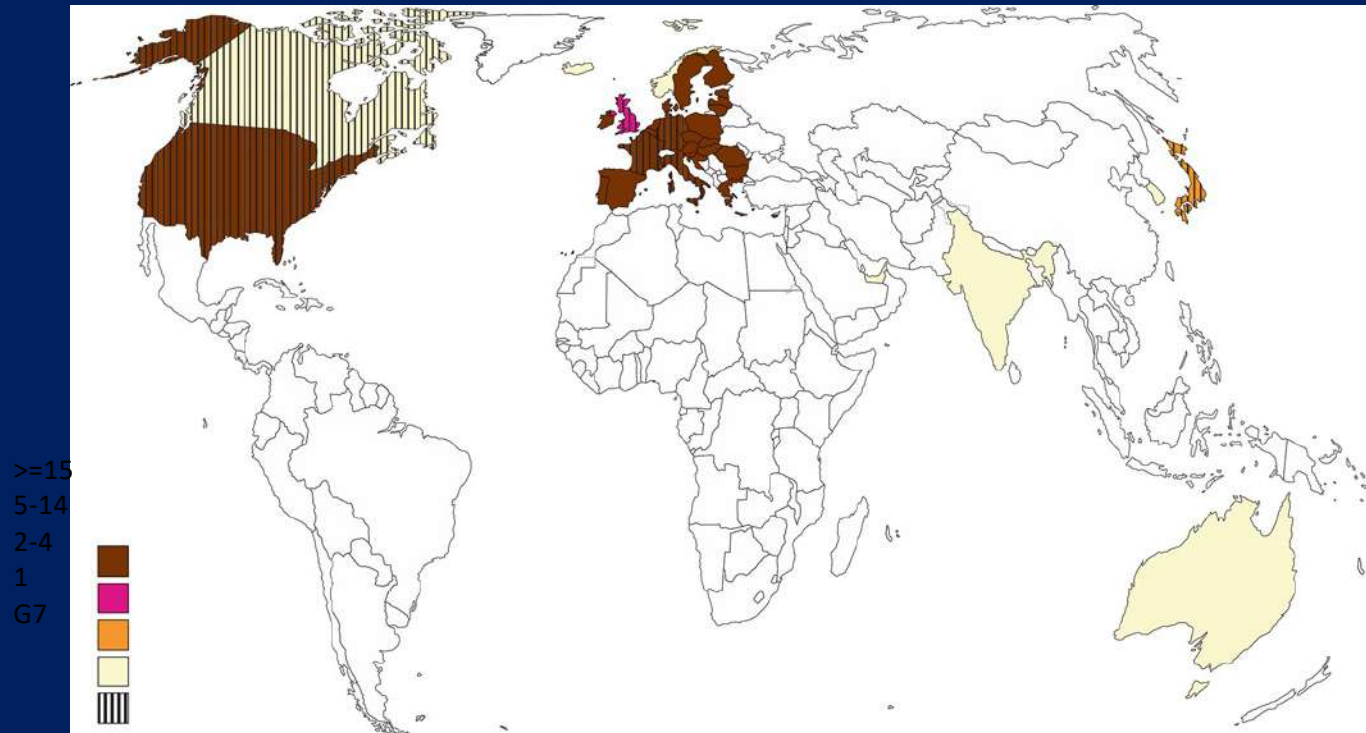
TRUSTWORTHY AI, EU VIEW



SOCIETAL AND ENVIRONMENTAL ASPECTS IN THE DESIGN OF AI



NUMBER OF CURRENT AI GUIDELINES, GLOBALLY



Geographic distribution of issuers of ethical AI guidelines by number of documents released. Most ethics guidelines are released in the United States ($n = 21$) and within the European Union (19), followed by the United Kingdom (13) and Japan (4). Canada, Iceland, Norway, the United Arab Emirates, India, Singapore, South Korea and Australia are represented with 1 document each. Having endorsed a distinct G7 statement, member states of the G7 countries are highlighted separately. <https://www.nature.com/articles/s42256-019-0088-2.pdf>

"AI READINESS"

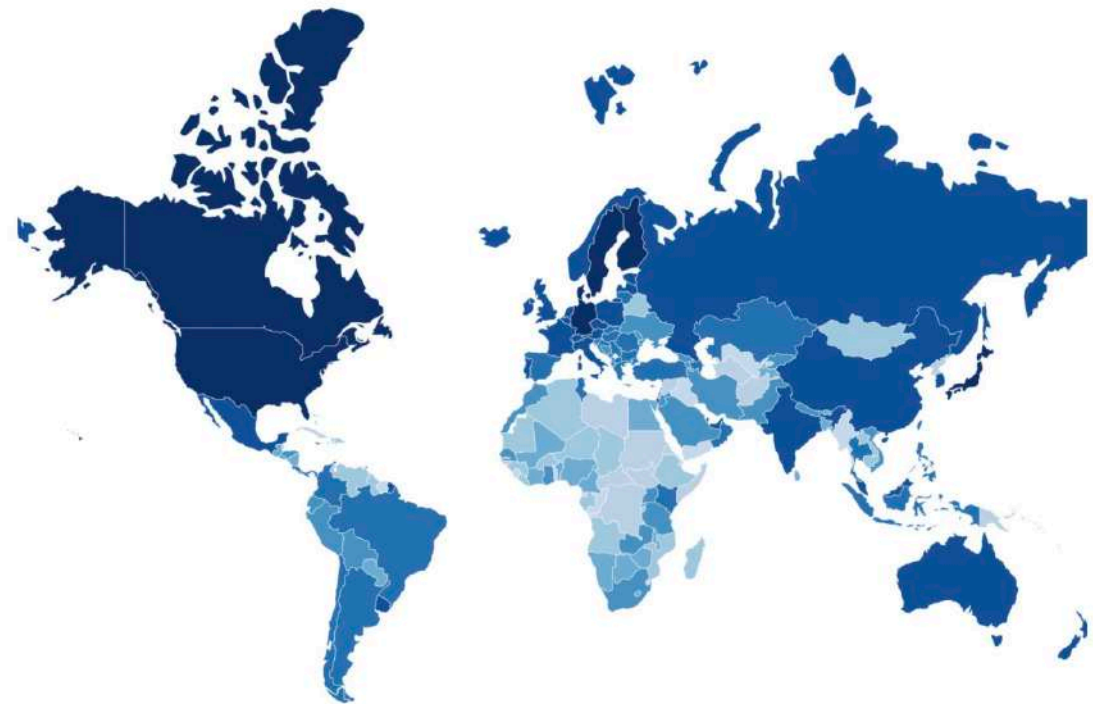
"Government AI Readiness Index" captures the current capacity of governments to exploit the innovative potential of AI.

The overall score is comprised of governance; infrastructure and data; skills and education; and government and public services.

REGULATION

State-Based Regulation
Industry-based Regulation
Citizen-based Regulation (crowdsourced)
Academic or Expert-Based Regulation

The Government AI Readiness Index 2019: How is Europe Doing in Comparison to the Rest of the World



The Government AI Readiness Index 2019

<https://www.humane-ai.eu/ai-readiness-index-2019/>

Ethics guidelines for AI by country of issuer, European

(Colors: Red = governmental + academia, Blue=companies, Black= groups)

Name of document/website	Issuer	Country
Position on Robotics and Artificial Intelligence	The Greens (Green Working Group Robots)	EU
Report with Recommendations to the Commission on Civil Law Rules on Robotics	European Parliament	EU
Ethics Guidelines for Trustworthy AI	High-Level Expert Group on Artificial Intelligence	EU
AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations	AI4People	EU
European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and Their Environment	Council of Europe: European Commission for the Efficiency of Justice (CEPEJ)	EU
Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems	European Commission, European Group on Ethics in Science and New Technologies	EU
Work in the Age of Artificial Intelligence. Four Perspectives on the Economy, Employment, Skills and Ethics	Ministry of Economic Affairs and Employment	Finland
Tieto's AI Ethics Guidelines	Tieto	Finland
Commitments and Principles	OP Group	Finland

Ethics guidelines for AI by country of issuer, European		
Name of document/website	Issuer	Country
How Can Humans Keep the Upper Hand? Report on the Ethical Matters Raised by AI Algorithms	French Data Protection Authority (CNIL)	France
For a Meaningful Artificial Intelligence. Towards a French and European Strategy	Mission Villani	France
Ethique de la Recherche en Robotique	CERNA (Allistene)	France
AI Guidelines	Deutsche Telekom	Germany
SAP's Guiding Principles for Artificial Intelligence	SAP	Germany
Automated and Connected Driving: Report	Federal Ministry of Transport and Digital Infrastructure, Ethics Commission	Germany
Ethics Policy	Icelandic Institute for Intelligent Machines (IIIM)	Iceland
L'intelligenza Artificiale al Servizio del Cittadino	Agenzia per l'Italia Digitale (AGID)	Italy
Human Rights in the Robot Age Report	The Rathenau Institute	Netherlands
Dutch Artificial Intelligence Manifesto	Special Interest Group on Artificial Intelligence (SIGAI), ICT Platform Netherlands (IPN)	Netherlands
Artificial Intelligence and Privacy	The Norwegian Data Protection Authority	Norway

Ethics guidelines for AI by country of issuer, Europe

Name of document/website	Issuer	Country
Principles of robotics	Engineering and Physical Sciences Research Council UK (EPSRC)	UK
The Ethics of Code: Developing AI for Business with Five Core Principles	Sage	UK
Big Data, Artificial Intelligence, Machine Learning and Data Protection	Information Commissioner's Office	UK
DeepMind Ethics & Society Principles	DeepMind Ethics & Society	UK
Business Ethics and Artificial Intelligence	Institute of Business Ethics	UK
AI in the UK: Ready, Willing and Able?	UK House of Lords, Select Committee on Artificial Intelligence	UK
Artificial Intelligence (AI) in Health	Royal College of Physicians	UK
Initial Code of Conduct for Data-Driven Health and Care Technology	UK Department of Health & Social Care	UK
Ethics Framework: Responsible AI	Machine Intelligence Garage Ethics Committee	UK
The Responsible AI Framework	PricewaterhouseCoopers UK	UK
Responsible AI and Robotics. An Ethical Framework.	Accenture UK	UK
Machine Learning: The Power and Promise of Computers that Learn by Example	The Royal Society	UK
Ethical, Social, and Political Challenges of Artificial Intelligence in Health	Future Advocacy	UK

Ethics guidelines for AI by issuer, International

Name of document/website	Issuer	Country
Artificial Intelligence and Machine Learning: Policy Paper	Internet Society	International
Report of COMEST on Robotics Ethics	COMEST/UNESCO	International
Ethical Principles for Artificial Intelligence and Data Analytics	Software & Information Industry Association (SIIA), Public Policy Division	International
ITI AI Policy Principles	Information Technology Industry Council (ITI)	International
Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2	Institute of Electrical and Electronics Engineers (IEEE), The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems	International
Top 10 Principles for Ethical Artificial Intelligence	UNI Global Union	International
The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation	Future of Humanity Institute; University of Oxford; Centre for the Study of Existential Risk; University of Cambridge; Center for a New American Security; Electronic Frontier Foundation; OpenAI	International
White Paper: How to Prevent Discriminatory Outcomes in Machine Learning	WEF, Global Future Council on Human Rights 2016-2018	International
Privacy and Freedom of Expression in the Age of Artificial Intelligence	Privacy International & Article 19	International

Ethics guidelines for AI by issuer, international

Name of document/website	Issuer	Country
The Toronto Declaration: Protecting the Right to Equality and Non-discrimination in Machine Learning Systems	Access Now; Amnesty International	International
Charlevoix Common Vision for the Future of Artificial Intelligence	Leaders of the G7	International
Artificial Intelligence: Open Questions About Gender Inclusion	W20	International
Declaration on Ethics and Data Protection in Artificial Intelligence	ICDPPC	International
Universal Guidelines for Artificial Intelligence	The Public Voice	International
Ethics of AI in Radiology: European and North American Multisociety Statement	American College of Radiology; European Society of Radiology; Radiology Society of North America; Society for Imaging Informatics in Medicine; European Society of Medical Imaging Informatics; Canadian Association of Radiologists; American Association of Physicists in Medicine	International
Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition (EAD1e)	Institute of Electrical and Electronics Engineers (IEEE), The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems	International

Ethics guidelines for AI by country of issuer, International

Name of document/website	Issuer	Country
Privacy and Freedom of Expression in the Age of Artificial Intelligence	Department of Industry Innovation and Science	Australia
Montréal Declaration: Responsible AI	Université de Montréal	Canada
Discussion Paper: National Strategy for Artificial Intelligence	National Institution for Transforming India (NITI Aayog)	India
The Japanese Society for Artificial Intelligence Ethical Guidelines	Japanese Society for Artificial Intelligence	Japan
Report on Artificial Intelligence and Human Society (unofficial translation)	Advisory Board on Artificial Intelligence and Human Society (initiative of the Minister of State for Science and Technology Policy)	Japan
Draft AI R&D Guidelines for International Discussions	Institute for Information and Communications Policy (IICP), The Conference toward AI Network Society	Japan
Sony Group AI Ethics Guidelines	Sony	Japan
Discussion Paper on Artificial Intelligence (AI) and Personal Data—Fostering Responsible Development and Adoption of AI	Personal Data Protection Commission Singapore	Singapore
Mid- to Long-Term Master Plan in Preparation for the Intelligent Information Society	Government of the Republic of Korea	South Korea
AI Principles of Telefónica	Telefónica	Spain
AI Principles & Ethics	Smart Dubai	UAE
Tenets	Partnership on AI	N/A
Principles for Accountable Algorithms and a Social Impact	Fairness, Accountability, and Transparency in Machine Learning	N/A

Ethics guidelines for AI by country of issuer, International

Name of document/website	Issuer	Country
Unified Ethical Frame for Big Data Analysis. IAF Big Data Ethics Initiative, Part A	The Information Accountability Foundation	USA
The AI Now Report. The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term	AI Now Institute	USA
Statement on Algorithmic Transparency and Accountability	Association for Computing Machinery (ACM)	USA
AI Principles	Future of Life Institute	USA
AI—Our Approach	Microsoft	USA
Artificial Intelligence. The Public Policy Opportunity	Intel Corporation	USA
IBM’s Principles for Trust and Transparency	IBM	USA
OpenAI Charter	OpenAI (a research laboratory based in San Francisco, California. investors include Microsoft, Reid Hoffman’s charitable foundation, and Khosla Ventures)	USA
Our Principles	Google	USA

Ethics guidelines for AI by country of issuer, International

Name of document/website	Issuer	Country
Policy Recommendations on Augmented Intelligence in Health Care H-480.940	American Medical Association (AMA)	USA
Everyday Ethics for Artificial Intelligence. A Practical Guide for Designers and Developers	IBM	USA
Governing Artificial Intelligence. Upholding Human Rights & Dignity	Data & Society	USA
Intel's AI Privacy Policy White Paper. Protecting Individuals' Privacy and Data in the Artificial Intelligence World	Intel Corporation	USA
Introducing Unity's Guiding Principles for Ethical AI—Unity Blog	Unity Technologies	USA
Digital Decisions	Center for Democracy & Technology	USA
Science, Law and Society (SLS) Initiative	The Future Society	USA
AI Now 2018 Report	AI Now Institute (at New York University)	USA
Responsible Bots: 10 Guidelines for Developers of Conversational AI	Microsoft	USA
Preparing for the Future of Artificial Intelligence	Executive Office of the President; National Science and Technology Council; Committee on Technology	USA
The National Artificial Intelligence Research and Development Strategic Plan	National Science and Technology Council; Networking and Information Technology Research and Development Subcommittee	USA
AI Now 2017 Report	AI Now Institute (at New York University - an interdisciplinary research center dedicated to understanding the social implications of artificial intelligence)	USA

ETHICAL PRINCIPLES IDENTIFIED IN EXISTING AI GUIDELINES

From: The global landscape of AI ethics guidelines (Anna Jobin, Marcello Lenca and Effy Vayena) <https://www.nature.com/articles/s42256-019-0088-2.pdf>

Ethical principle	Number of documents	Included values
Transparency	73/84	Transparency, explainability, explicability, understandability, interpretability, communication, disclosure
Justice and fairness	68/84	Justice, fairness, consistency, inclusion, equality, equity, (non-) bias, (non-)discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access and distribution
Non-maleficence	60/84	Non-maleficence, security, safety, harm protection, precaution, prevention, integrity (bodily or mental), non-subversion
Responsibility	60/84	Responsibility, accountability, liability, acting with integrity
Privacy	47/84	Privacy, personal or private information, Integrity
Beneficence	41/84	Benefits, beneficence, well-being, peace, social good, common good
Freedom and autonomy	34/84	Freedom, autonomy, consent, choice, self-determination, liberty, empowerment
Trust	28/84	Trust
Sustainability*	14/84	Sustainability, environment (nature), energy, resources (energy)
Dignity	13/84	Dignity
Solidarity	6/84	Solidarity, social security, cohesion

*<https://www.aisustainability.org/> AI SUSTAINABILITY CENTER Stockholm

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>