CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG

AI FROM PHILOSOPHY TO ETHICS TO SCIENCE TO TECHNOLOGY TO LAW AND BACK. Thoughts from an European perspective

GORDANA DODIG-CRNKOVIC

CHALMERS UNIVERSITY OF TECHNOLOGY

GOTHENBURG, SWEDEN

CHAIR AI Ethics Seminar Series

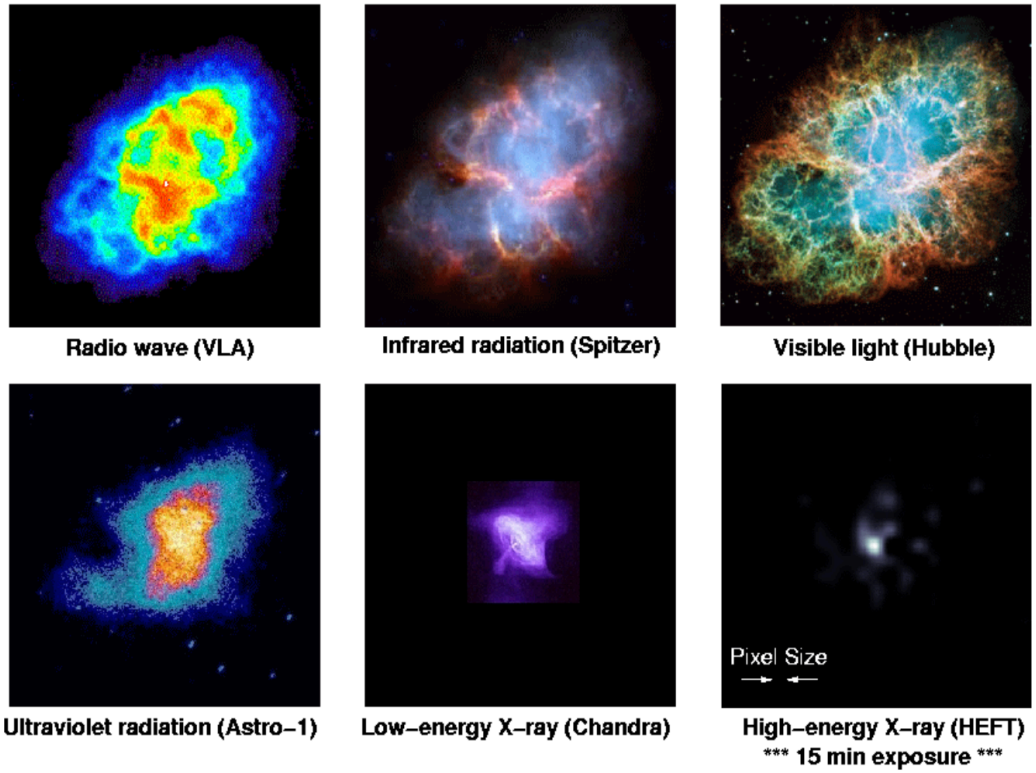
Gothenburg, 25th February 2020

http://gordana.se/Presentations

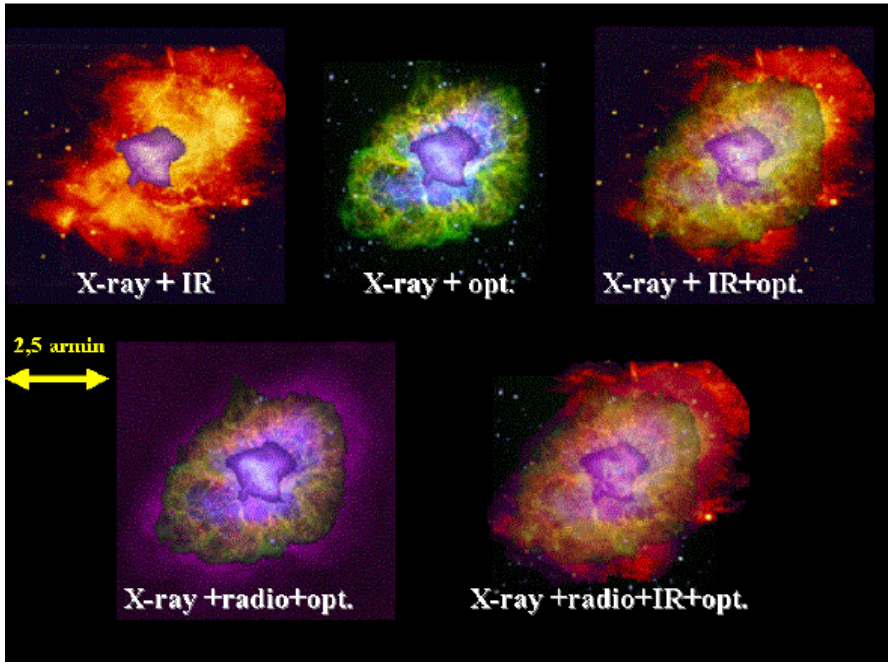
https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/AI genie in a bottle

THE SAME OBJECT, IN DIFFERENT WAVELENGTHS

Crab Nebula: Remnant of an Exploded Star (Supernova)



Superposition of wavelengths



CHALMERS
UNIVERSITY OF TECHNOLOGY

UNIVERSITY OF GOTHENBURG

TRANSDISCIPLINARY APPROACH:

WE WILL MAKE A SUPERPOSITION OF DIFFERENT VIEWS IN ORDER TO GET A MORE COMPLETE UNDERSTANDING OF AI ETHICS IN ITS CONTEXT

My background: theoretical physics, computer science,
philosophy of computing and ethics,
computational models of cognition,
recent interest: interaction design
interdisciplinary/transdisciplinary elucidation of AI.

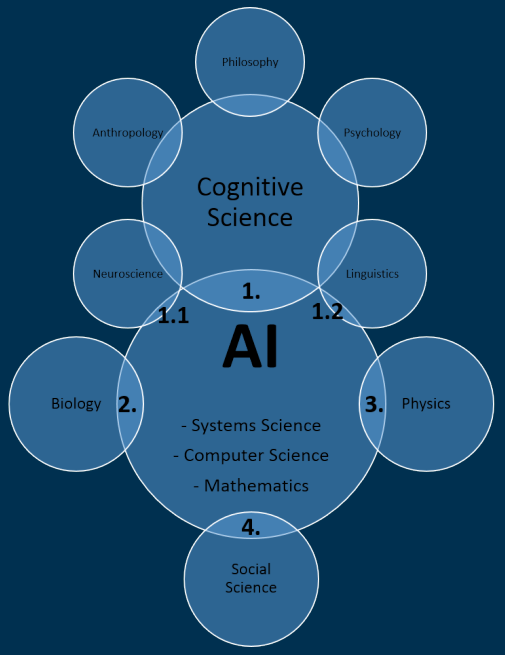
PHILOSOPHICAL ASPECTS OF AI

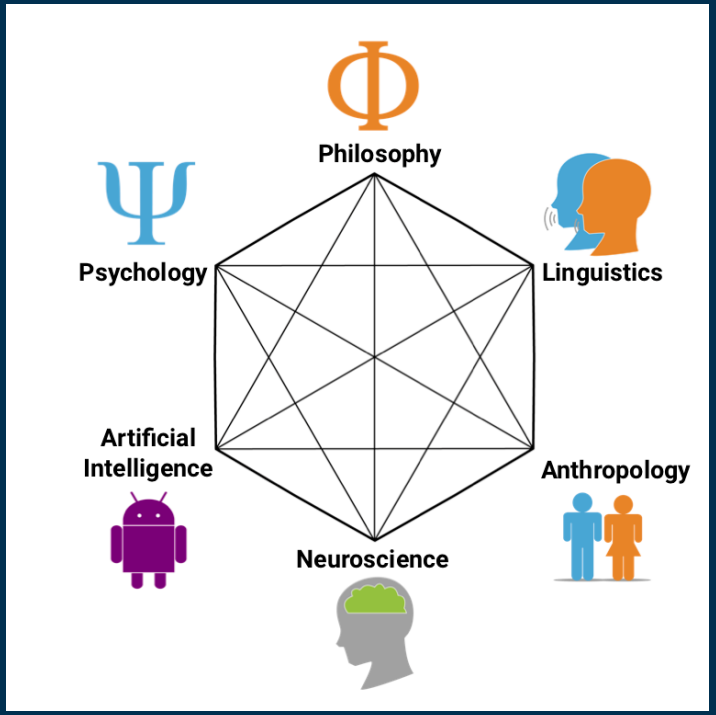
Philosophy → Metaphysics → Ontology
What is AI? (AI that exists and can exist)

Philosophy → Epistemology
Knowledge of AI
(What and how we know about AI)

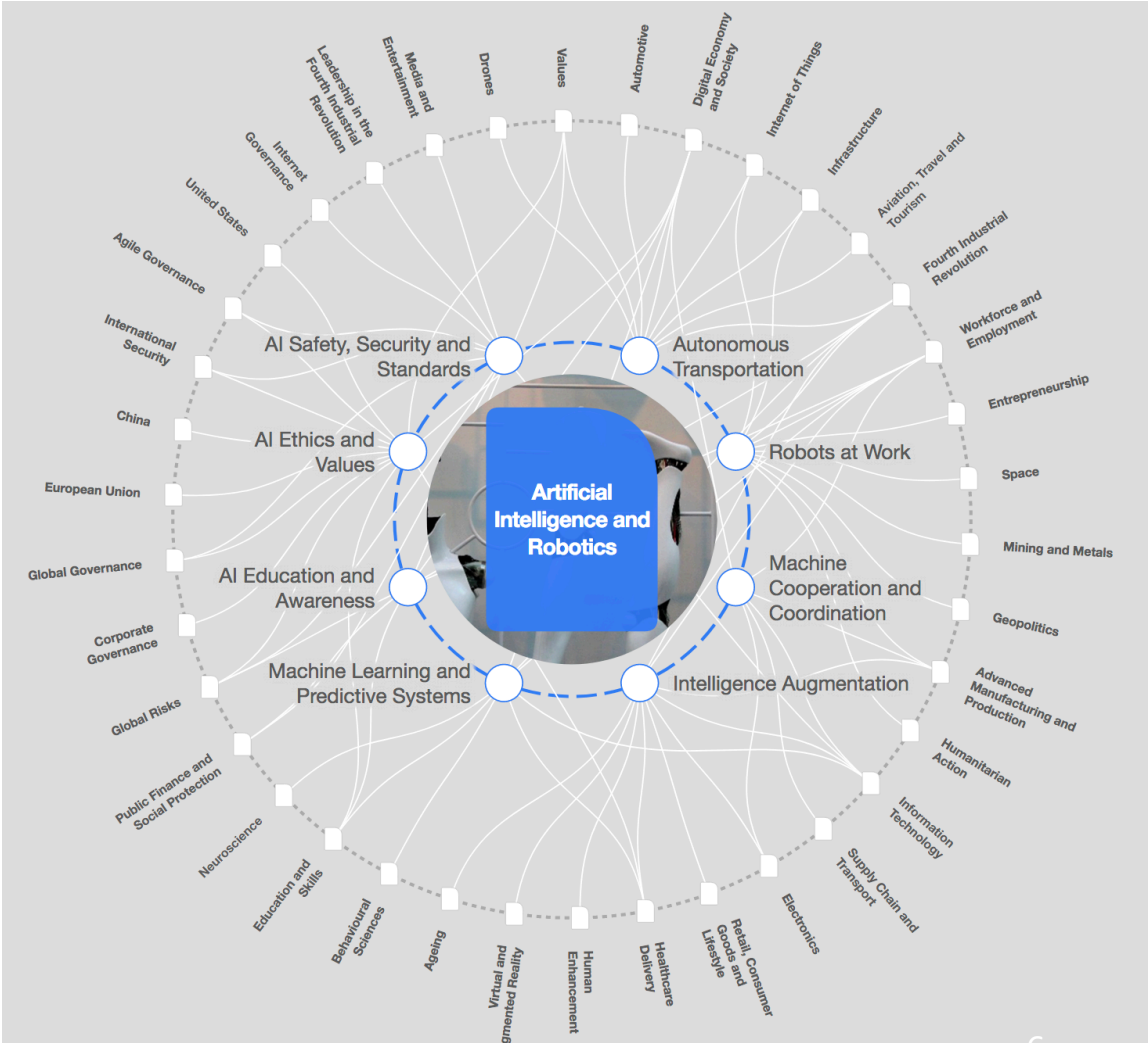
Philosophy → Axiology → Ethics
Good and Bad
Right and Wrong in AI

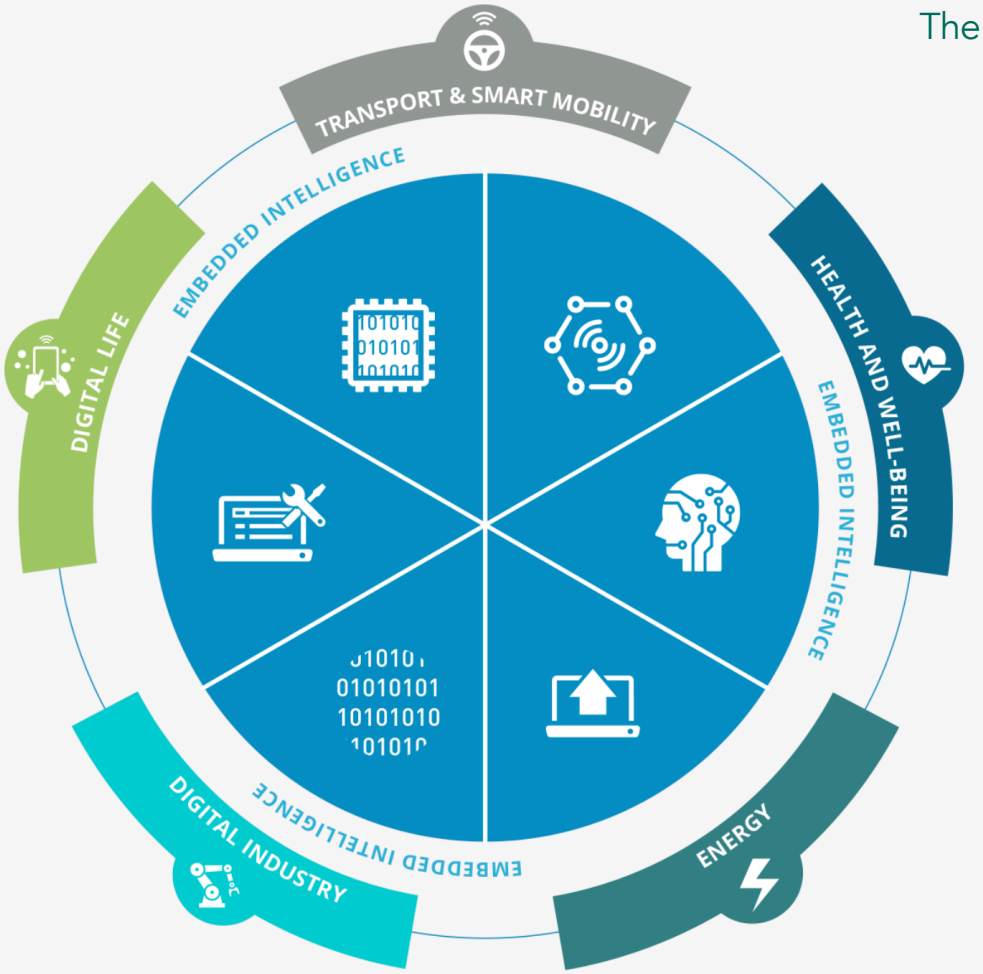
SCIENTIFIC ASPECTS OF AI





AI AND TECHNOLOGY





The six technologies that create Embedded Intelligence

- Embedded and Cyber-Physical Systems
- Secure IoT and System of Systems
- EdgeComputing & Embedded Artificial Intelligence
- Embedded High-Performance Computing
- System of Systems Integration platforms for Digitalisation
- Embedded Software Technologies and Software engineering Tools

https://artemis-ia.eu

LEGAL ASPECTS 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

Mirelle Hilderbrandt

- 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

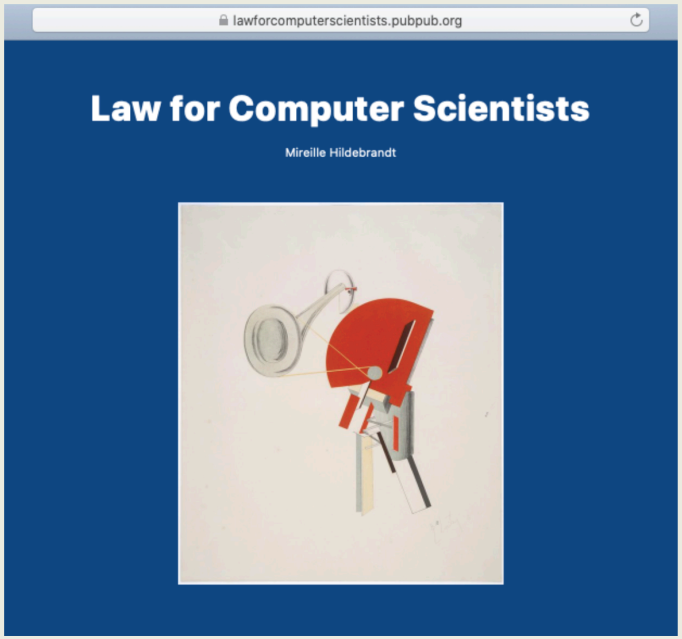
Available at MIT's pubpub:

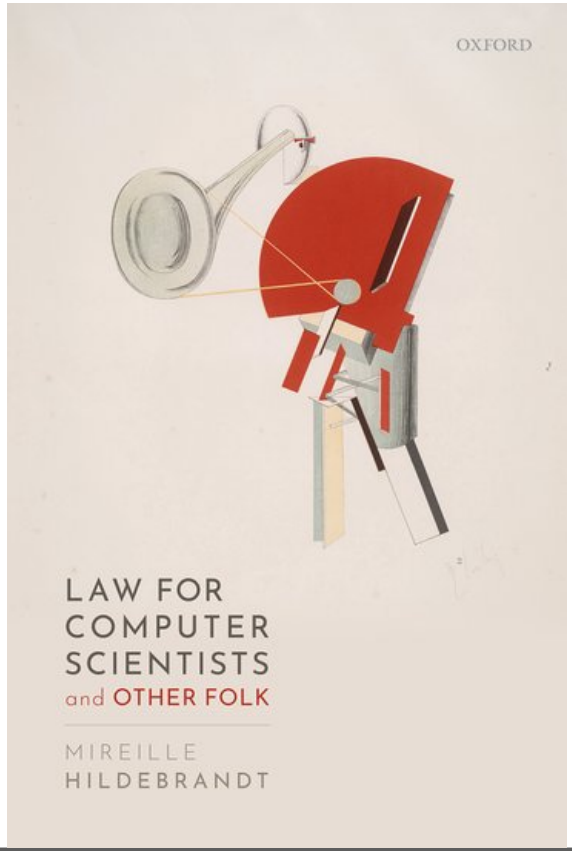
- https://lawforcomputerscientists.pubpub.org

In print March 2020

Oxford University Press

- Hardcopy
- Ebook in open access





LAW FOR COMPUTER SCIENTISTS

Mirelle Hilderbrandt

https://lawforcomputerscientists.pubpub.org/

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**.

ANCIENT ROOTS OF AI

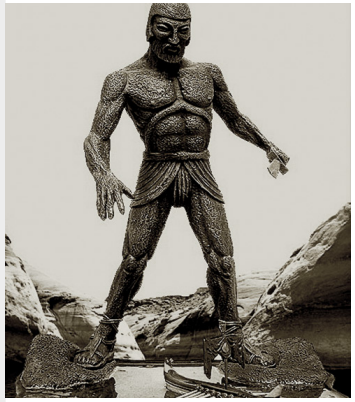
TALOS OF CRETE-

FIRST INTELLIGENT ROBOT & THE ANTIKYTHERA MECHANISM

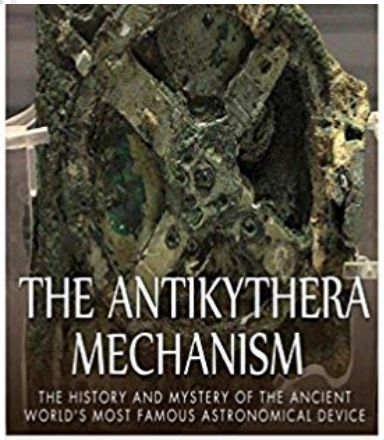
ANCIENT ANALOG COMPUTER

Talos was a mythical bronze age (3200 to 1200 B.C.E.) giant, the first robot in history, which protected Minoan Crete from invaders. Talos was not born but made, either by Zeus himself or, according to other versions of the myth, by the Hephaestus, god of fire and iron, on Zeus's order.

Antikythera was 1st century BC analog computer, designed to calculate astronomical positions.



https://taloscrafts.com/who-is-talos/



https://en.wikipedia.org/wiki/Antikythera_mechanism

INTELLIGENT ARTIFACTS TODAY

WITH DIFFERENT INTELLIGENT PROPERTIES

- Ambient intelligence
- Intelligent robots & softbots
- Intelligent transportation systems
- Intelligent cities, Intelligent IoT
- Decision making algorithms
- AI for health
- Scientific AI
- AI for software, etc.

Leading companies

Amazon

Apple

Facebook

Google

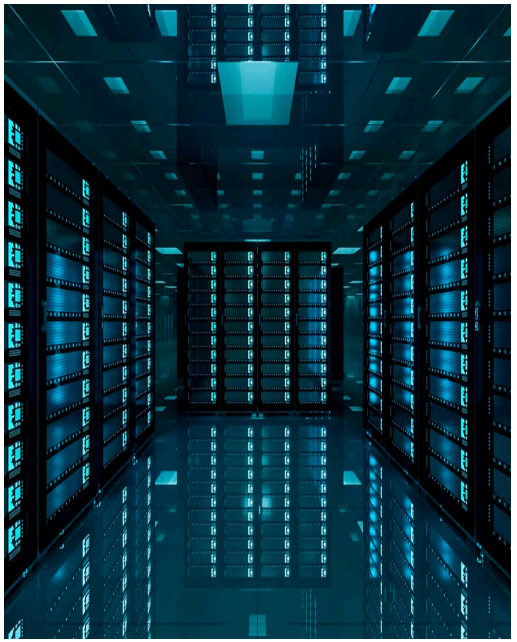
IBM

Intel

Microsoft

Nvidia

Twitter



https://bitcoinist.com/crypto-mining-becoming-concern-us-cities/

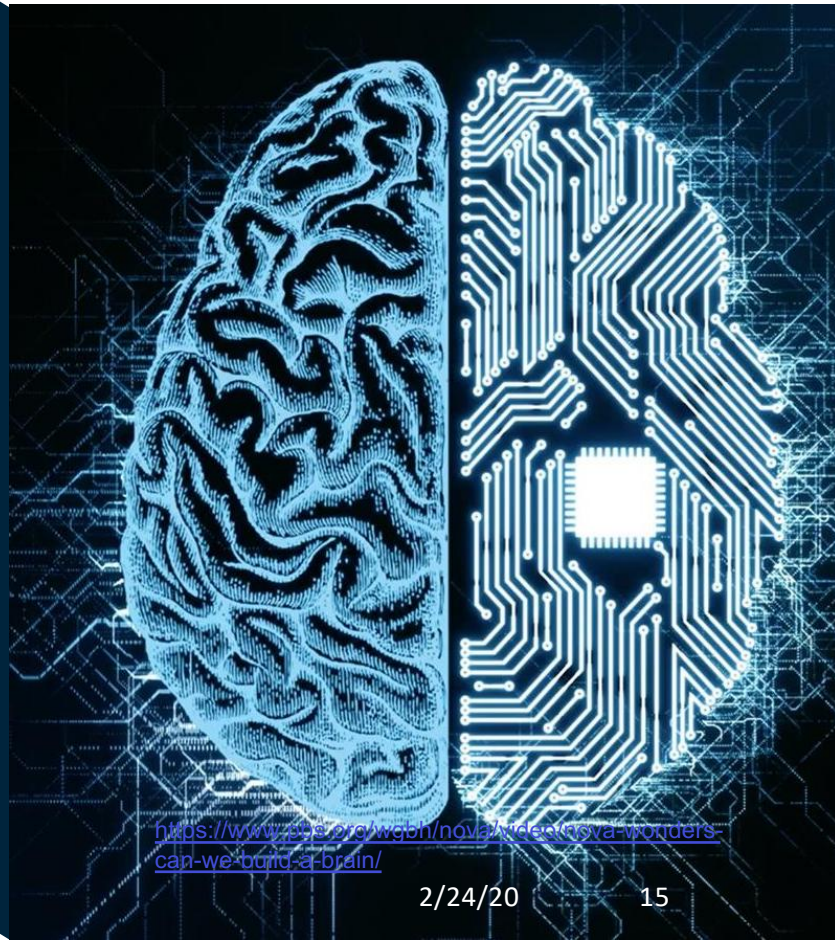
AI & HUMAN INTELLIGENCE: AI AS ENHANCER

Cognitive enhancements. Restoring & enhancing memory. Creating artificial memories.

Theodore Berger (University of Southern California, L.A.) Engineering Memories: A Cognitive Neural Prosthesis for Restoring and Enhancing Memory Function https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3141091/ https://www.technologyreview.com/s/513681/memory-implants/

Developing a hippocampal neural prosthetic to facilitate human memory encoding and recall, Robert E Hampson et al 2018 J. Neural Eng. 15 036014 https://iopscience.iop.org/article/10.1088/1741-2552/aaaed7/pdf

A Successful Artificial Memory Has Been Created. Scientific American https://tiny.cc/obuxbz The growing science of memory manipulation raises social and ethical questions. By Robert Martone on August 27, 2019



https://www.pbs.org/wgbh/nova/video/nova-wonders-can-we-build-a-brain/

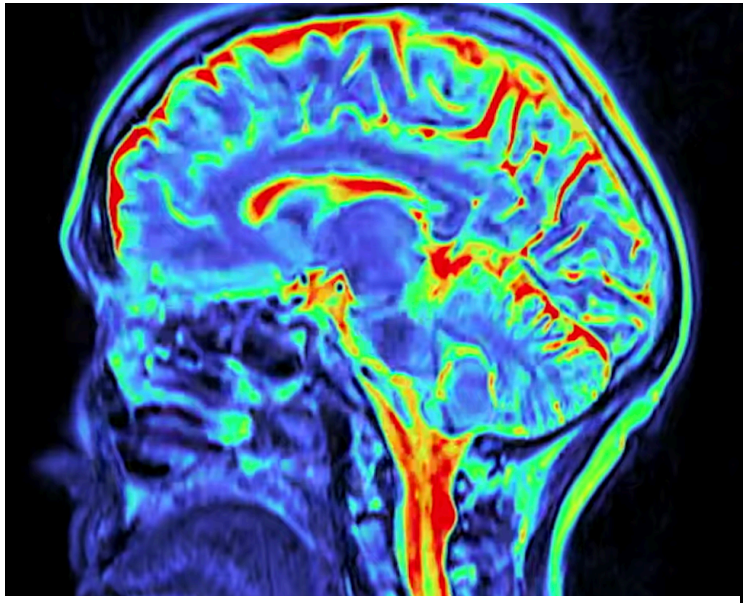
AI & HUMAN INTELLIGENCE
AI AS REPLACER

"Mind reading" technologies. MIT, The University of California in San Francisco, Elon Mask and Facebook are "in the race to read minds with computers"- in order to enable merging the human brain with the computer
https://www.youtube.com/watch?v=R3G5fzz76lQ

AI mimicking human perception, reasoning and planning

Decision-making intelligent programs

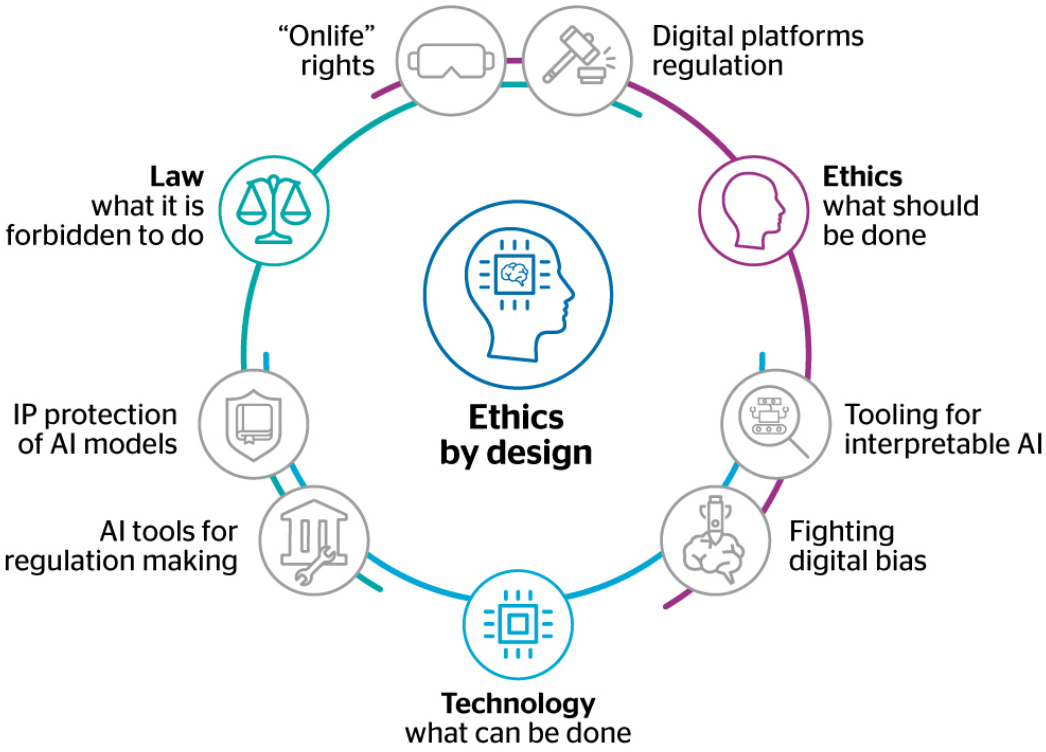
AI can be autonomous, unpredictable, biased and opaque/inscrutable unlike classical technologies



DO WE WANT INTELLIGENT EVERYTHING?

DO WE WANT TO **CONTROL** EVERYTHING?

WHY? QUESTION OF VALUES & ETHICS



https://atos.net/content/mini-sites/journey-2022/human-centric-ai/

VALUE-SENSITIVE DESIGN

Value-sensitive design (VSD) is based on the insight that artefacts are value-laden and design is value-sensitive. We need to identify early implicit values embedded in new technologies by focusing on the use of technology.

"Value" is defined broadly as property that a person or a group considers important in life, and designers can intentionally or unintentionally inscribe their values in the design objects thus shaping them.

The design is carried out iteratively by combining the following approaches supporting the values:

- conceptual (conceptions of values for users and stakeholders),
- empirical (how values are realized in practice)
- technical (design of technology),
- research all of which is followed by
- assessment

ETHICAL AI DESIGN EXPECTATIONS

- EXPLAINABLE & ACCOUNTABLE AI
- PROMOTING HUMAN RIGHTS
- RESPECTING GDPR (PROTECTING PRIVACY, PERSONAL INTEGRITY)
- FAIR, TRANSDPARENT, ACCOUNTABLE SYSTEMS
- SAFETY CRITICAL AI MUST BE REGULATED, CERTIFIED, WITH REGULATORY OVERSIGHT

TRANSPARENT ETHICAL GOVERNANCE

- DEMOCRATIC LEGITIMACY

- CURRENT CONCENTRATION OF POWER IN THE HANDS OF THE FEW (GOOGLE, APPLE, AMAZON, MICROSOFT)

- FOLLOWING RESPONSIBLE RESEARCH & INNOVATION (EUROPE)*

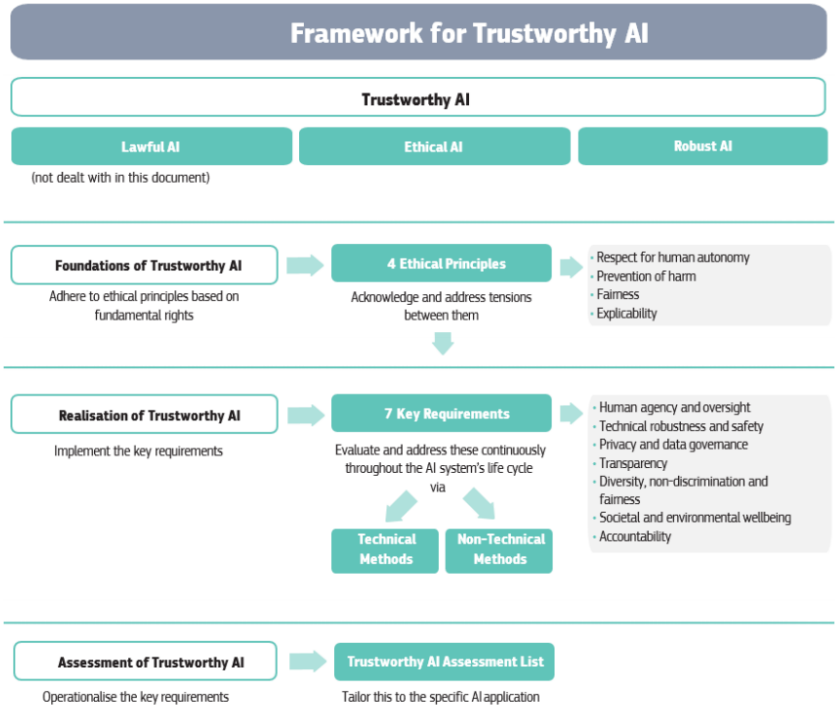
- IEEE INITIATIVE – EDUCATION – ETHICS IN THE WHOLE PROCESS OF ENGINEERING

- NEW TYPES OF PROBLEMS IN “AUTOMATED PUBLIC SPHERE”

- STAKEHOLDERS INVOLVEMENT IS ESSENTIAL

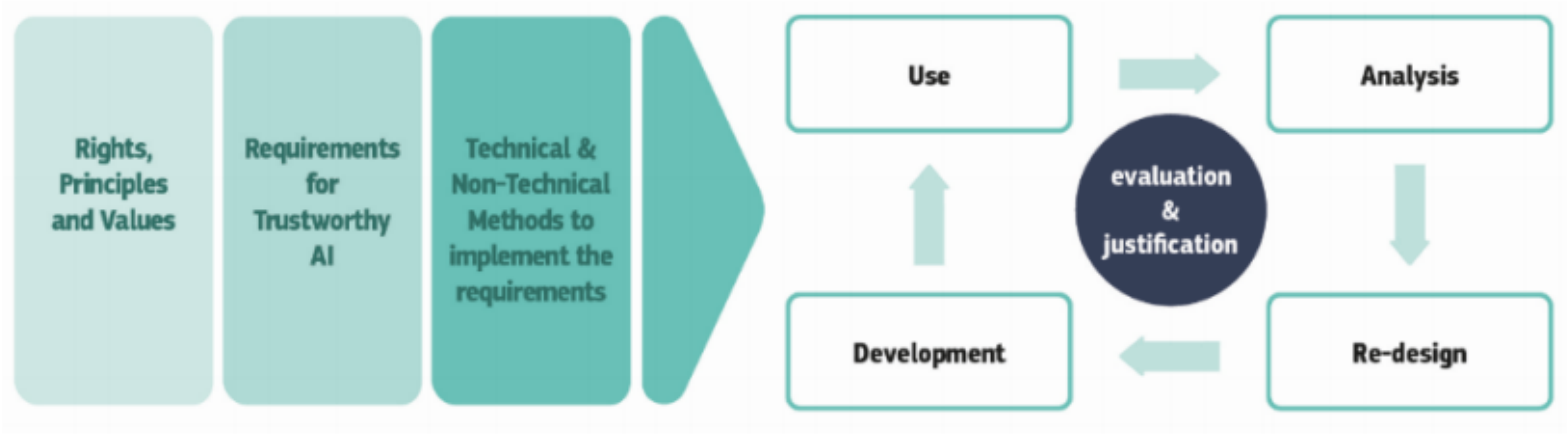
*https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation

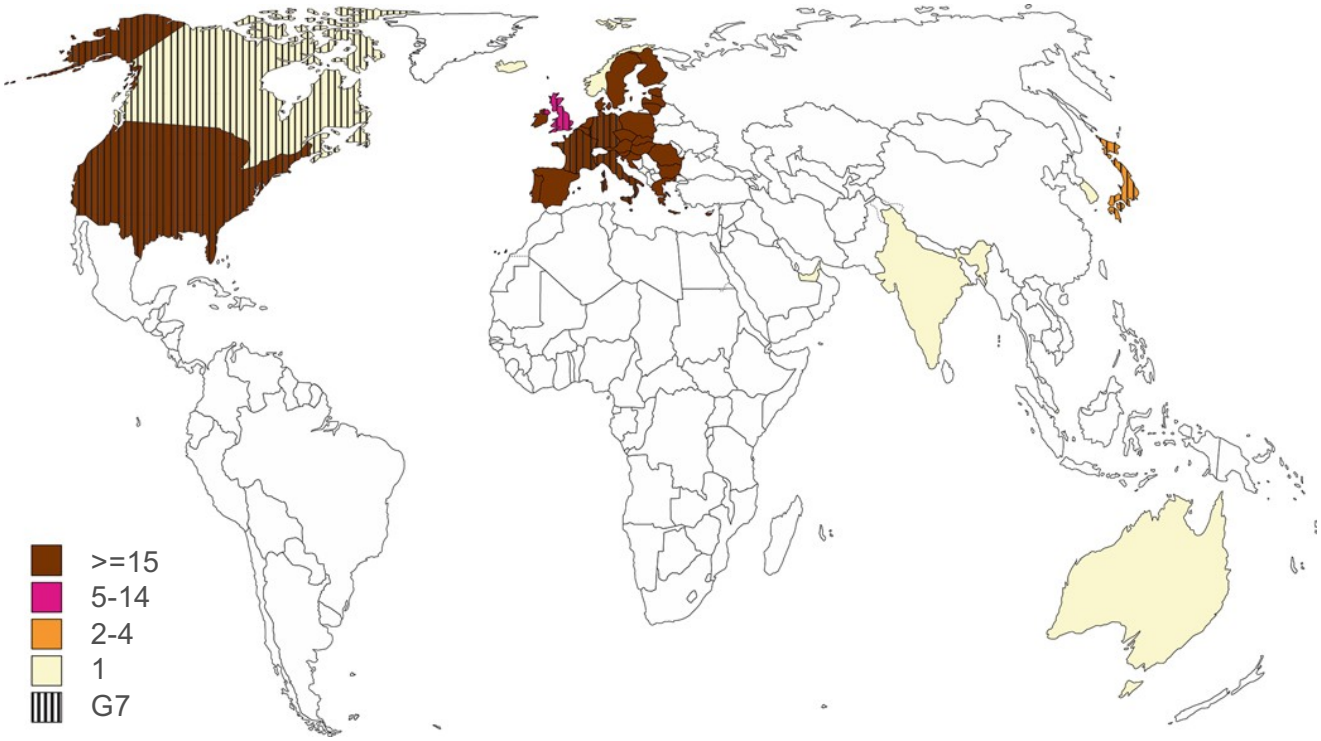
TRUSTWORTHY AI, EU VIEW



Ethics Guidelines for Trustworthy AI, EU Commission

SOCIETAL AND ENVIRONMENTAL ASPECTS





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

May 21, 2019

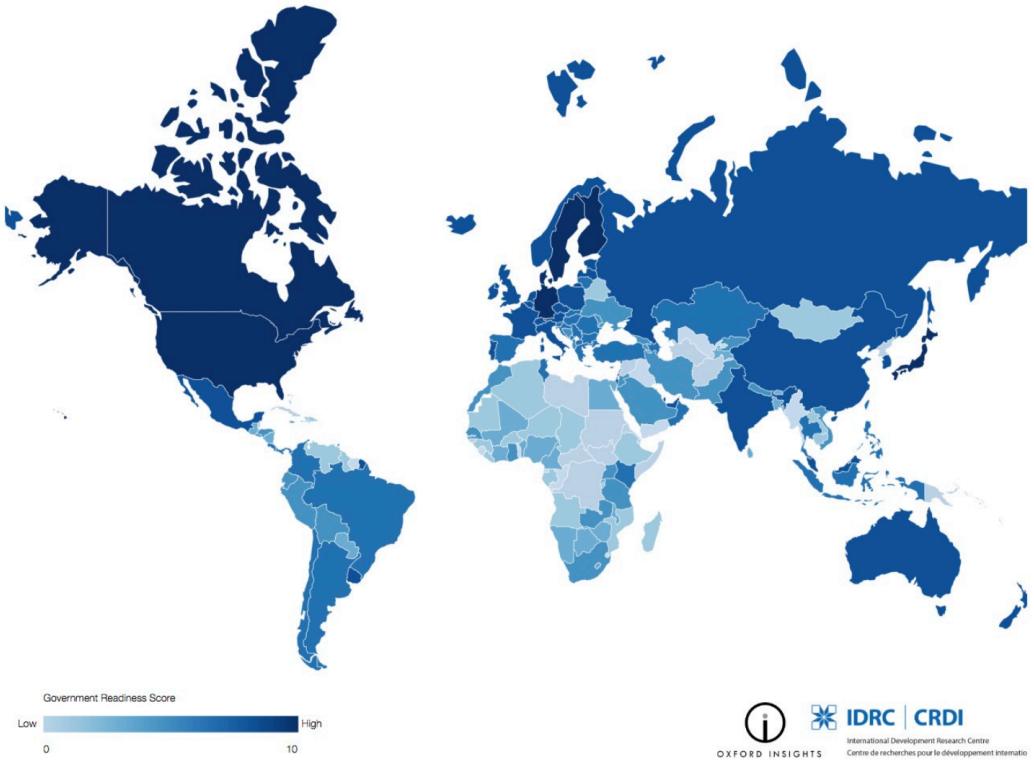
News

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

AI readiness

In 2017, Oxford Insights created the world's first Government AI Readiness Index, to answer the question: how well placed are national governments to take advantage of the benefits of AI in their operations and delivery of public services? The results sought to capture the current capacity of governments to exploit the innovative potential of AI.

The overall score is comprised of 11 input metrics, grouped under four high-level clusters: governance; infrastructure and data; skills and education; and government and public services. The data is derived from a variety of resources, ranging from our own desk research into AI strategies, to databases such as the number of registered AI startups on Crunchbase, to indices such as the UN eGovernment Development Index.



The Government AI Readiness Index 2019

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

Ethics guidelines for AI by country of issuer

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

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'intelligenzia 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

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 country of issuer

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 country of issuer

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

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 Statement for Algorithms	Fairness, Accountability, and Transparency in Machine Learning (FATML)	N/A
10 Principles of Responsible AI	Women Leading in AI	N/A

Ethics guidelines for AI by country of issuer

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

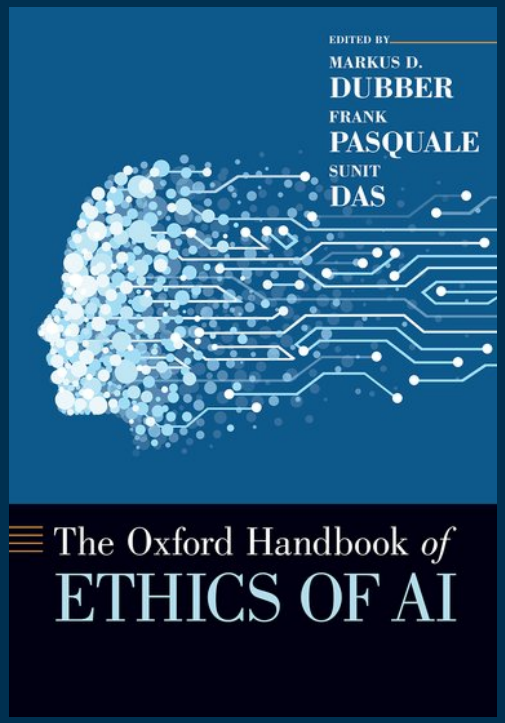
Ethics guidelines for AI by country of issuer

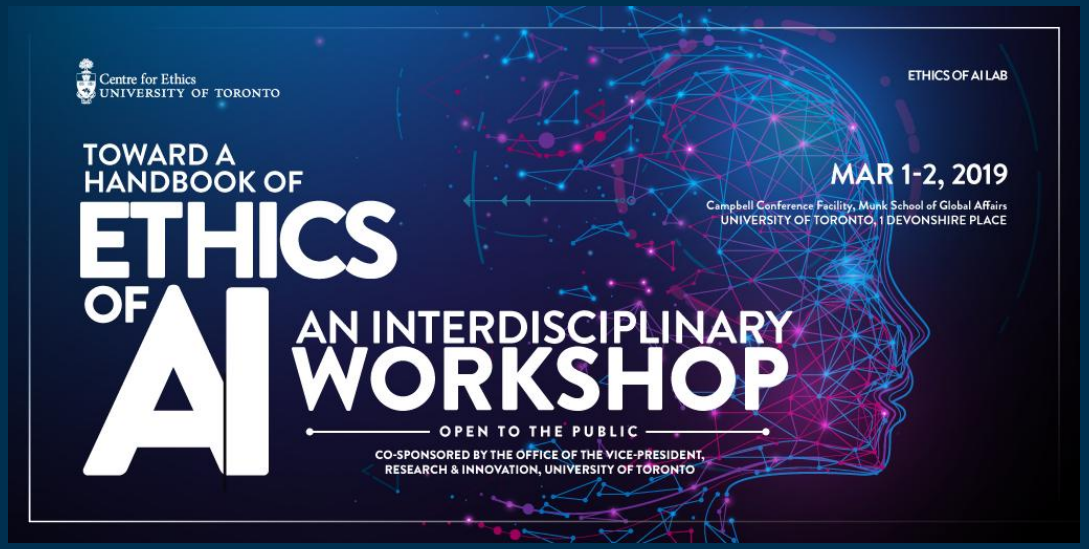
Name of document/website	Issuer	country
Policy Recommendations on Augmented Intelligence in Health Care H-480.940	American Medical Association (AMA)	USA
Everyday Ethics for Artificial Intelligence. A Practical Guide for Designers and Developers	IBM	USA
Governing Artificial Intelligence. Upholding Human 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 - an interdisciplinary research center dedicated to understanding the social implications of artificial intelligence)	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	USA

ETHICAL PRINCIPLES IDENTIFIED IN EXISTING AI GUIDELINES

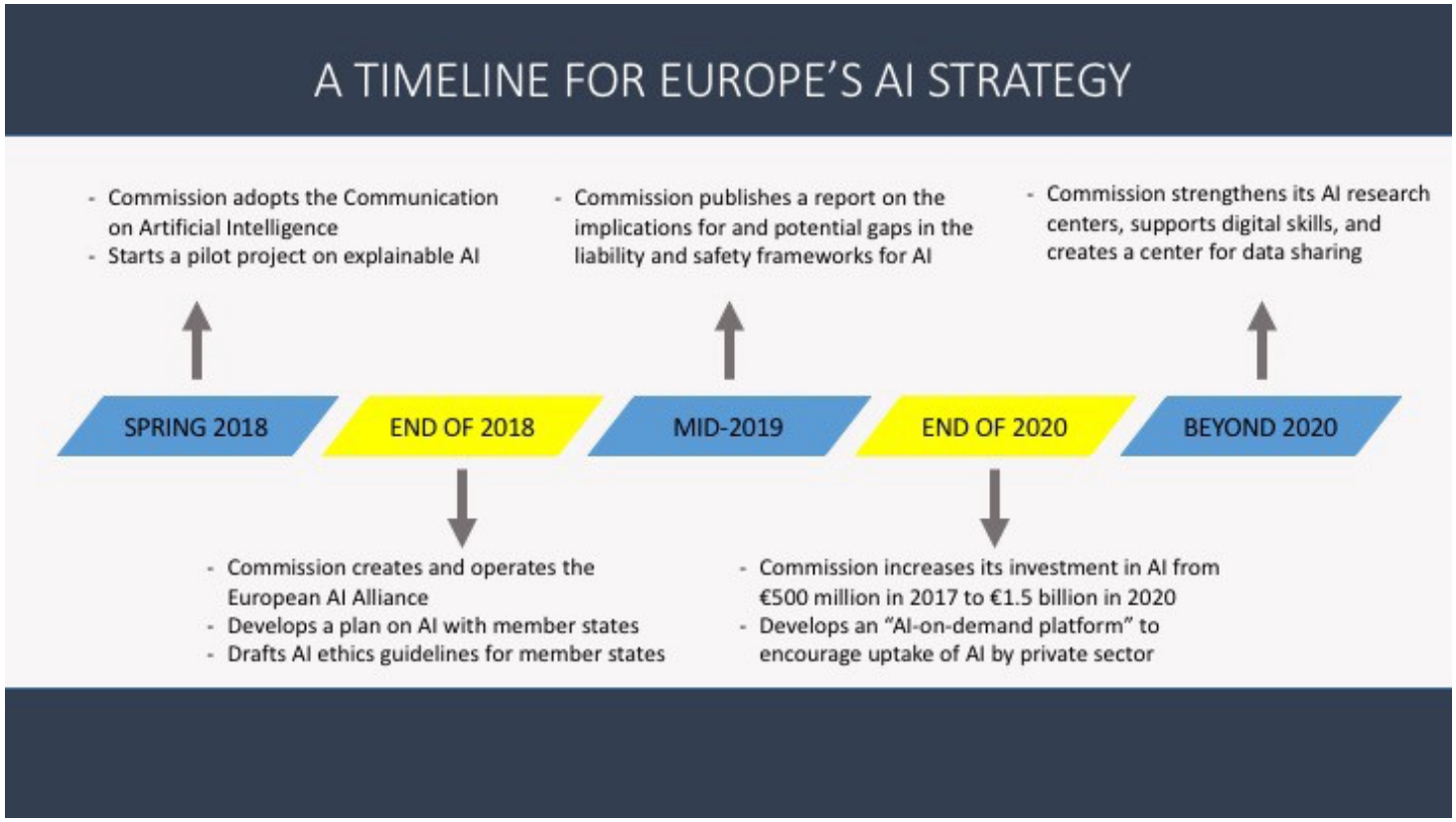
https://www.nature.com/articles/s42256-019-0088-2.pdf

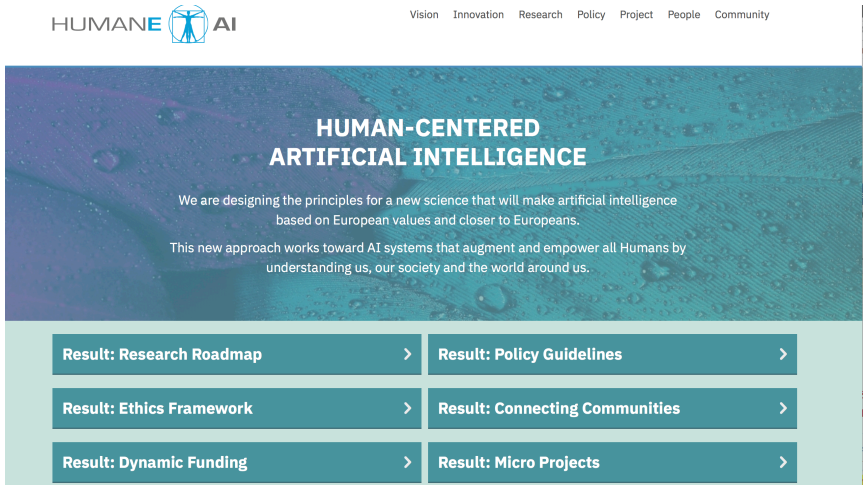
Ethical principle	Number of documents	Included values
Transparency	73/84	Transparency, explainability, explicability, understandability, interpretability, communication, disclosure
Justice and fairness	68/84	Justice, fairness, consistency, inclusion, equality, equity, (non-) bias, (non-)discrimination, 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, autonomy, 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



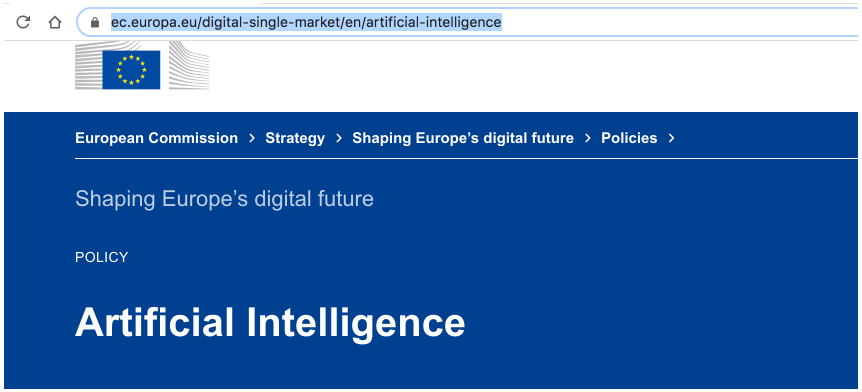


https://global.oup.com/academic/product/the-oxford-handbook-of-ethics-of-ai-9780190067397?cc=ca&lang=en&#


A TIMELINE FOR EUROPE'S AI STRATEGY
- Commission adopts the Communication on Artificial Intelligence
- Starts a pilot project on explainable AI
- Commission publishes a report on the implications for and potential gaps in the liability and safety frameworks for AI
- Commission strengthens its AI research centers, supports digital skills, and creates a center for data sharing
SPRING 2018
END OF 2018
MID-2019
END OF 2020
BEYOND 2020
- Commission creates and operates the European AI Alliance
- Develops a plan on AI with member states
- Drafts AI ethics guidelines for member states
- Commission increases its investment in AI from €500 million in 2017 to €1.5 billion in 2020
- Develops an "AI-on-demand platform" to encourage uptake of AI by private sector



https://www.humane-ai.eu



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

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

http://www.ai4people.eu

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

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

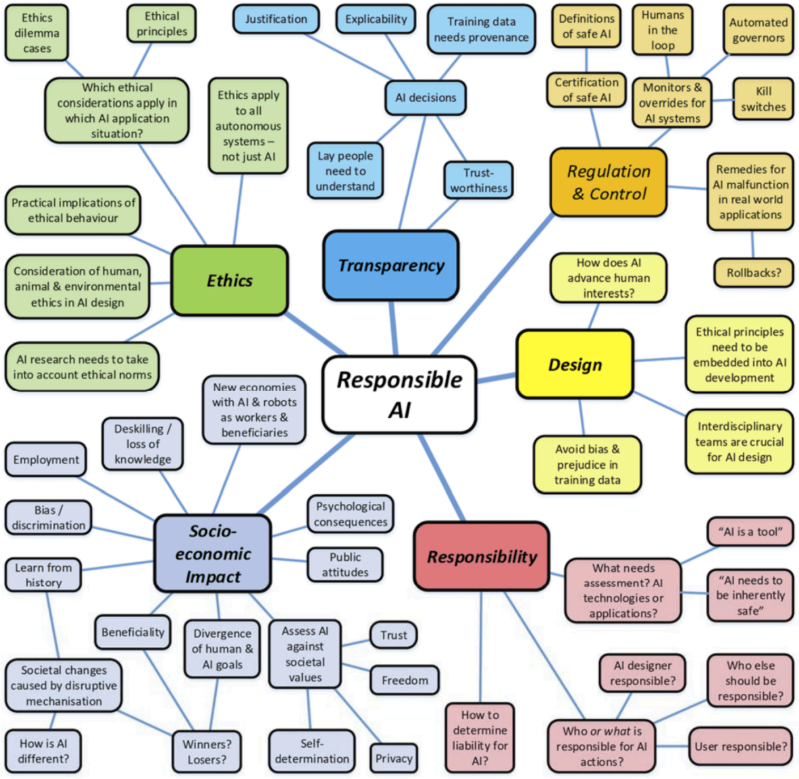
Dr. Matthias Leese, ETH Zurich, CH

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

CHALMERS UNIVERSITY OF TECHNOLOGY | UNIVERSITY OF GOTHENBURG

WHITE PAPER ON AI

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

First analysis of the EU Whitepaper on AI

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

Whitepaper correctly promotes the need for 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.

EUROPEAN
COMMISSION

Brussels, 19.2.2020
COM(2020) 65 final

WHITE PAPER

On Artificial Intelligence - A European approach to excellence and trust

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

President Margarethe Vestager explained to a group of journalists recently, the Whitepaper basically says, in very legal language: "let's pause and figure out if there are any situations, and if so, under what circumstances facial recognition should be authorised". And she added: "as it stands right now, GDPR would say 'don't use it', because you cannot get consent."

https://theintercept.com/2020/02/21/eu-facial-recognition-database/

Leaked Reports Show EU Police Are Planning a Pan-European Network of Facial Recognition Databases

AI in the Flow- Learning AI! From Philosophy to Ethics to Science to Technology to Law 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 of AI, followed tightly of technological applications comes.

Laws 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. Friedman, B., Kahn, P.H., Jr. (2003) Human values, ethics, and design. In The Human-Computer Interaction Handbook, Fundamentals, Evolving Technologies and Emerging Applications; Jacko, J.A., Sears, A., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA; pp. 1177–1201.

2. Friedman, B., Kahn, P.H., Jr., Borning, A. (2006) Value sensitive design and information systems. In Human-Computer Interaction in Management Information Systems: Applications; M.E. Sharpe, Inc.: New York, NY, USA; Volume 6, pp. 348–372.

3. IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems. Ethically Aligned Design, Version One – For Public Discussion (2016) A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v1.pdf

4. IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design, Version 2 for Public Discussion (2017) A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems. Available online: https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead_v2.pdf

5. Asilomar Conference 2017. Asilomar AI Principles. Available online: https://futureoflife.org/ai-principles/?cn-reloaded=1

6. 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

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. https://www.aisustainability.org/ AI SUSTAINABILITY CENTER Stockholm

7. https://ai4good.org/ AI FOR GOOD

8. http://gordana.se/

REFERENCES

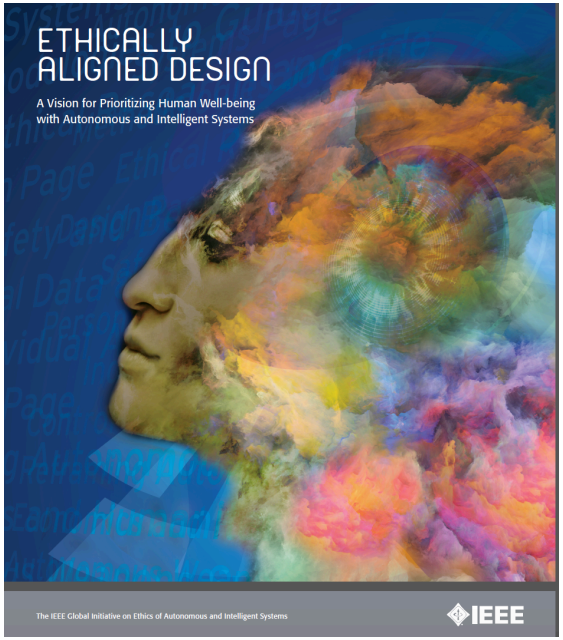
6. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F.; et al. (2018) AI4People—An Ethical Framework for a Good AI Society. Minds Mach. 28, 689–707. https://link.springer.com/article/10.1007%2Fs11023-018-9482-5

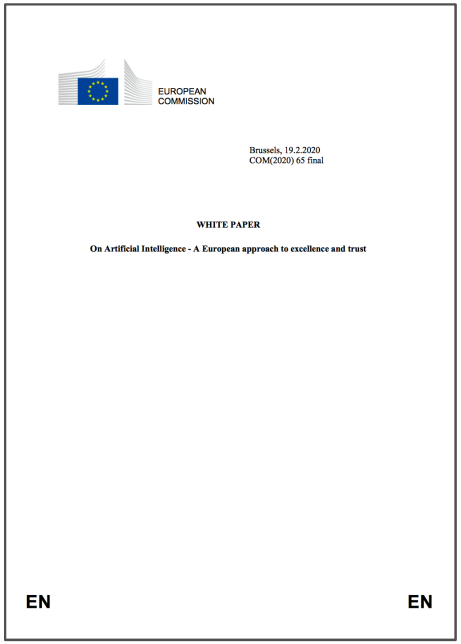
7. Floridi, L. (2019) Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. Philosophy & Technology. https://doi.org/10.1007/s13347-019-00354-x

8. 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

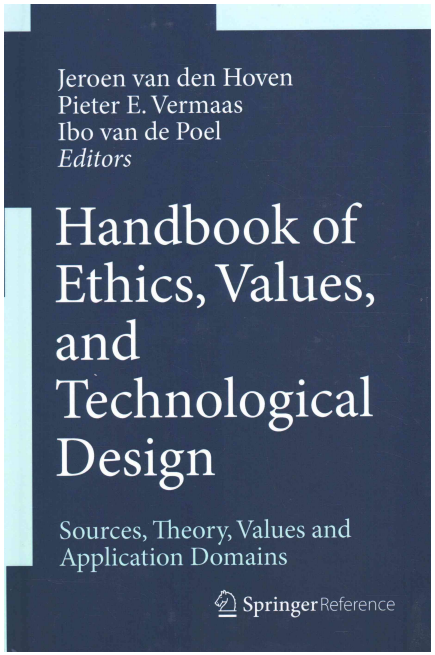
9. Wachter S, Mittelstadt B, Floridi L (2017) Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, vol. 7, issue 2 (2017) pp. 76-99 Published by Oxford University Press (OUP)

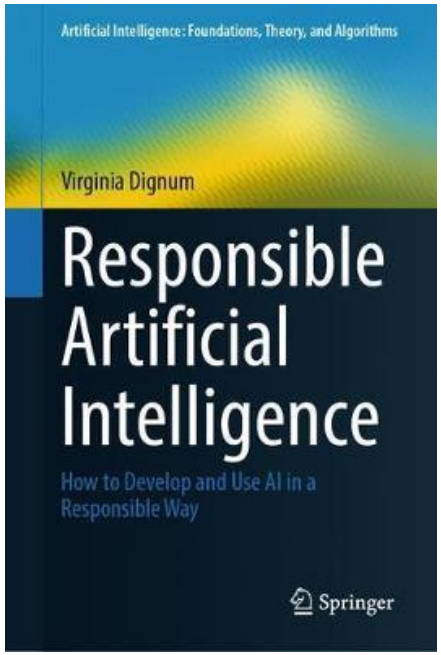
10. Dodig-Crnkovic G. and Çürüklü B., Robots - Ethical by Design, Ethics and Information Technology 2011, Volume 14, Number 1, pp. 61-71. http://www.springerlink.com/content/f432g33181787u63/fulltext.html



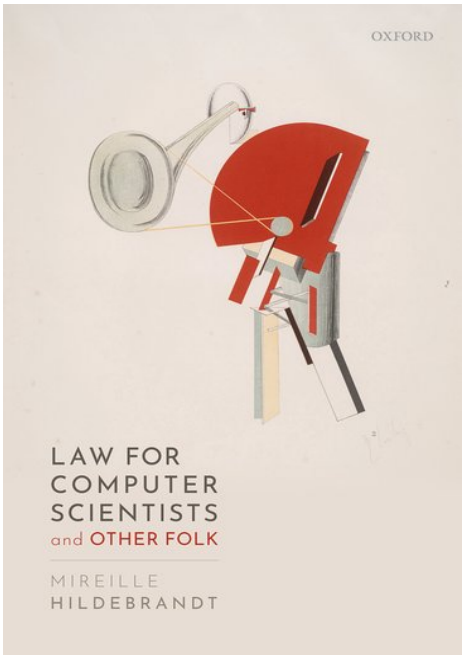


https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai


Jeroen van den Hoven
Pieter E. Vermaas
Ibo van de Poel
Editors
Handbook of Ethics, Values, and Technological Design
Sources, Theory, Values and Application Domains
SpringerReference


Artificial Intelligence: Foundations, Theory, and Algorithms
Virginia Dignum
Responsible Artificial Intelligence
How to Develop and Use AI in a Responsible Way
Springer


Sarah Spiekermann
Ethical IT Innovation
A Value-Based System Design Approach


OXFORD
LAW FOR COMPUTER SCIENTISTS and OTHER FOLK
MIREILLE HILDEBRANDT

ETHICALLY ALIGNED DESIGN STANDARDS

The IEEE P7000™ series of standards projects under development addresses specific issues at the intersection of technological and ethical considerations. Like its technical standards counterparts, the IEEE P7000 series empowers innovation across borders and enables societal benefit.

The IEEE P7000™ - IEEE Standards Project Model Process for Addressing Ethical Concerns During System DesignInspired by Methodologies to Guide Ethical Research and Design Committee, and supported by IEEE Computer Society https:// standards.ieee.org/project/7000.html

IEEE P7001™ - IEEE Standards Project for Transparency of Autonomous SystemsInspired by the General PrinciplesCommittee, and supported by IEEE Vehicular Technology Society https://standards.ieee.org/project/7001.html

IEEE P7002™ - IEEE Standards Project for Data Privacy Process Inspired by The Personal Data and Individual Agency Control Committee, and supported by IEEE Computer Society https://standards.ieee.org/project/7002.html

IEEE P7003™ - IEEE Standards Project for Algorithmic Bias ConsiderationsSupported by IEEE Computer Society https://standards.ieee.org/project/7003.html

•IEEE P7004™ - IEEE Standards Project for Child and Student Data GovernanceInspired by The Personal Data and Individual Agency Control Committee, and supported by IEEE Computer Society https://standards.ieee.org/project/7004.htm

All links accessed on 15 August 2019

CHALMERS
UNIVERSITY OF TECHNOLOGY

|

UNIVERSITY OF GOTHENBURG

http://gordana.se