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

PART 2
ARTIFICIAL INTELLIGENCE, TECHNOLOGY AND VALUES

GORDANA DODIG-CRNKOVIC
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 part of the lecture covered the following:

Part 1 philosophical and transdisciplinary scientific view

Now we proceed to the technological overview and value-based
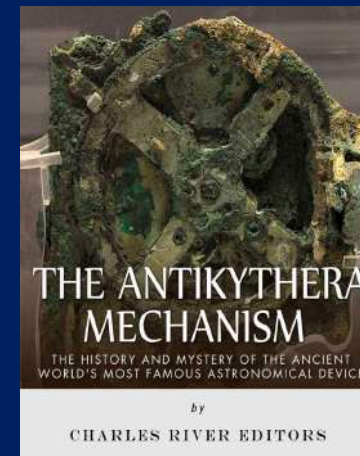foundations of AI

ANCIENT ROOTS OF AI IDEA

TALOS OF CRETE- FIRST INTELLIGENT ROBOT & THE ANTIKYTHERA MECHANISM ANCIENT ANALOG COMPUTER

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

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

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 of stars and planets.

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/

APPLICATIONS OF AI IN DIFFERENT DOMAINS
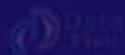
EMBEDDED INTELLIGENCE TECHNOLOGIES
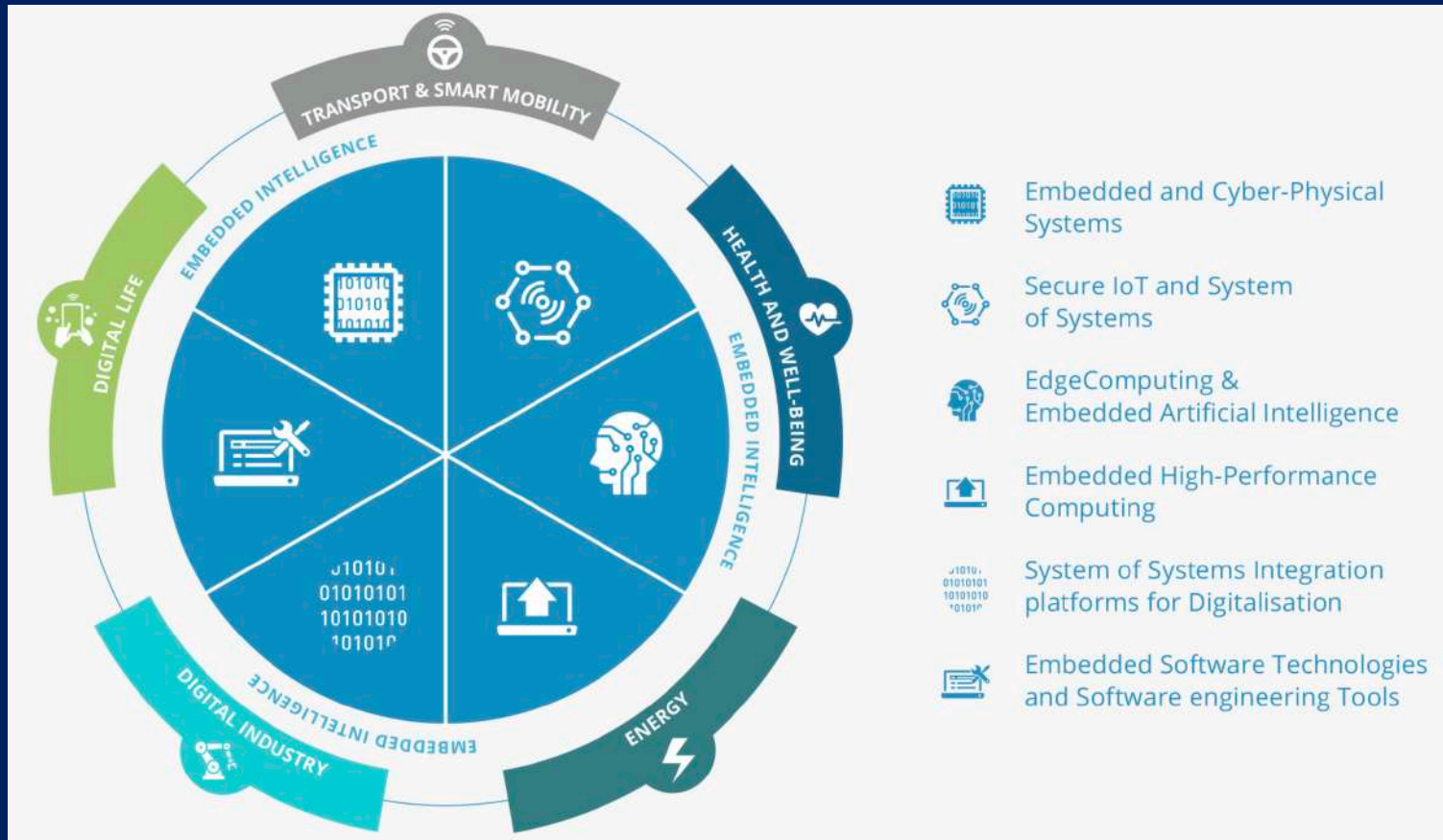
ALGORITHMIC AUTHORITY

Authority is increasingly expressed algorithmically according to the book "The Black Box Society,"* by Frank Pasquale (2016)

Because of information overload is the reason why we depend on algorithms to *sort, categorize, and prioritize* which information we will read and in which way.

Our reality is increasingly shaped by algorithms** . Power of platforms (contrary to the illusion of "technical neutrality and progressive openness")

Algorithmic biases affecting personal integrity (invisible discrimination driven by particular interests and social power)

Algorithmic culture - Humans defined by algorithms

* https://www.hup.harvard.edu/catalog.php?isbn=9780674970847
** David Beer (2017) The social power of algorithms, Information, Communication & Society, 20:1, 1-13, DOI: 10.1080/1369118X.2016.1216147
https://doi.org/10.1080/1369118X.2016.1216147

ALGORITHMIC AUTHORITY

Example: Netflix company that shapes the *media industry with recommendation engine- shaping both audience habits*

Algorithmic authority: algorithmically curating news and social media feeds, evaluating job performance, matching dates, algorithmic management – hiring, supervision, firing employees. (e.g., Facebook's friend feed and Google's search algorithms, Uber's algorithms and Amazon Mechanical Turk.

Algorithms are increasingly used for social, economical and political governance.***

***Katzenbach, C. & Ulbricht, L. (2019). Algorithmic governance. Internet Policy Review, 8(4). DOI: 10.14763/2019.4.1424. https://policyreview.info/node/1424/pdf

AI ETHICS & VALUE-SENSITIVE DESIGN

To connect ethics with AI technology, we need to be able to decide what is right and good in technology, its use and its effect on individuals, society and environment.

We have seen examples of intelligent technologies where AI has important ethical consequences such as *algorithmic governance* or *algorithmic decision-making.*

As we learned in the first part of this lecture, ethics is a branch of Philosophy dealing with the question "what is right?" or "what is good?"

As technology provides us with artifacts that are result of the process of design, our expectation is that they are intended for good of people.

AI ETHICS & VALUE-SENSITIVE DESIGN

One way to introduce ethics into the AI is through Value-sensitive design which is based on the insight that artefacts are value-loaded. We identify values embedded in technologies by studying its use.

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

For AI technology some of important values are: *safety, security, privacy, autonomy, trust, fairness, non-maleficence, beneficence, reliability, responsibility, sustainability*.

The design is typically carried out iteratively and values are being assimilated by combining the following approaches : conceptual, technical – empirical and research, with a continuous assessment and learning process within the ecology of socio-technological system.

ETHICAL AI DESIGN - EXPECTATIONS

- EXPLAINABLE & ACCOUNTABLE AI

- PROMOTING HUMAN RIGHTS

- PROTECTING PRIVACY, PERSONAL INTEGRITY (RESPECTING GDPR*)

- FAIR, TRANSDPARENT, ACCOUNTABLE SYSTEMS

- SAFETY CRITICAL AI MUST BE REGULATED, CERTIFIED,

  WITH REGULATORY OVERSIGHT

*General Data Protection Regulation

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. Spiekermann S. (2015) Ethical IT Innovation: A Value-Based System Design Approach. Taylor & Francis

6. Lustig, C., Pine, K., Nardi, B., Irani, L., Lee, M. K., Nafus, D., & Sandvig, C. (2016). Algorithmic authority: The ethics, politics, and economics of algorithms that interpret, decide, and manage. In CHI EA 2016: #chi4good - Extended Abstracts, 34th Annual CHI Conference on Human Factors in Computing Systems (Vol. 07-12-May-2016, pp. 1057-1062). Association for Computing Machinery. https://doi.org/10.1145/2851581.2886426

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 March 2020