Nr. 8 · March 2020

The Danish Board of Technology Foundation

A 360-degree evaluation of AI: Recommendations for action

WHY SHOULD WE DISCUSS AI AND ITS IMPLICATIONS?

Easy access to vast amounts of data, fast data processing, pattern recognition and sophisticated learning algorithms are tools that herald incredible possibilities for effectiveness, accuracy, abundance and potentially revolutionary changes to society.

As mind blowing as the possibility for improvement in almost all conceivable aspects of private and public life is, as worrying is it that AI technologies also have obvious capacities for abuse of power, for overruling existing norms and agreements regarding rights and integrities, and for downright malign and adverse application. The question is: Can we adapt our uptake of technology in the present to influence its direction and impact in the future? And what are the actions we can and should undertake?

In the AI 360 COPENHAGEN workshop we analysed the possible future impacts of AI across ethics, human rights, economy, society and policy to come up with recommendations for action.

Facts box:

- **Setting:** AI 360 Copenhagen was a multi-criteria expert workshop, drawing on the *AI 360 facilitation tool* (360tool.eu) to provide a multi-criteria analysis of AI.

- **Work:** The participants' deliberations on the future implications of AI was organised in five 'dimensions': Political implications, Rights and Ethics, Legal framework, Economy, and Societal implications.

- **Conclusions:** The good news: we may act to ensure the positive impacts of AI on our societies. The less good news: No good solutions exist to counter expected unequal distribution of cost and benefits related to the implementation of AI, AI also has a great potential for abuse.

This newsletter is a condensed version of the AI 360 project report, available at: wwww.humanbrainproject.eu

https://sos-ch-dk-2.exo.io/public-website-production/filer_public/a4/f2/a4f2aabd-6821-4a8d-b082-5070e2797b27/ai360_humanbrainproject_recommendations_report_final.pdf

Ai 360° copenhagen

March 21, 2019 11:30 am → March 22, 2019 3:00 pm
Strandgade 6, st. 1401 Copenhagen K

No good recommendations for dealing with questions on inequality and abuse

We brought together a number of European experts to discuss the uncertainties and challenges artificial intelligence poses to us. We asked them to give an overview of the political, social, ethical, legal, rights related and economic implications of artificial intelligence. We also asked them to point out what solutions could be used to deal with the challenges, and to assess how good a solution would be for dealing with the uncertainties they identified.

However, even if the recommendations from the experts were to be followed red flags remained. The experts did not see any good solutions for dealing with inequality in the distribution of benefits and risks that would follow from an increased implementation of AI in our societies, neither in the short or long term. In the short term (2025), social cohesion and inclusion also remain major red flags for which no adequate solutions exist at the present time.

In addition to the red flags, concerns remained on the potential of AI for abusive applications affecting fundamental rights and freedoms and the functioning of democratic societies.

The biggest challenge we face with artificial intelligence is the potential for abuse. Abuse that can take many different forms. In the political space, we have already seen how social media can be used to deceive, create distrust, influence referendums, and short-circuit democratic processes by, for example, spreading false allegations and bypassing the official office.

In a future of increased use of, and access to, artificial intelligence, the risk of this type of abuse will increase significantly. At the same time, we are facing a situation where global companies control the digital channels that are increasingly used for communication channels (e.g. Facebook, Twitter and Google), that are beyond democratic and legal control in many of the individual countries and territories where the platforms are used. The key question here is what role artificial intelligence will play in generating and sharing information in the future, and who is going to control that development? There are no sure answers to both questions right now.

According to the experts, trust in public authorities, the legal system, our politicians and the governments, is largely based on transparency. Understanding that it is possible to follow the process leading to, for example, a legal decision. Artificial intelligence systems challenge transparency as it may become impossible to follow the process behind an analysis or decision.

There are two related issues that come into play. First, transparency is dependent on information being shared and the information being shared is correct, but who decides what is correct? Which leads to the second challenge of introducing bias into an analysis process. It is impossible to design a technical solution without some kind of bias, so which bias should be introduced in a system, who should decide it and how can we overlook the consequences of the introduced bias? How do we ensure justice in, for example, our judicial system, whose decisions about, for example, the removal of children are made based on the use of systems based on artificial intelligence?

It fundamentally challenges the trust and coherence of our society if transparency becomes less. In addition, there is a risk of increased inequality between those who have access to resources that will enable them to challenge analysis and decisions made by an 'intelligent' system.

For example, there is a potential to simulate societal developments and trends and use such analysis to allocate resources in the health care system and to reduce misuse of public funds. The big challenge, with artificial intelligence is that it can do no more than the data it is fed. And the discriminatory patterns are something that people have created.

Ethical guidelines are a step on the path to increase reflection on built-in bias and unintended implications of new technological tools. In addition, education, quality assurance of artificial intelligence, and increased protection of individuals and their personal data are important steps towards realizing the positive gains of artificial intelligence. Unfortunately, they are not enough to solve basic societal imbalances in the distribution of power or resources that are only likely to be amplified in the future.

POLITICAL IMPLICATIONS

In their discussion on the political implications of AI, the experts elaborated on issues of: Transparency and trust; political culture and distribution of power; and equality and fairness.

Transparency and trust

- **Non-transparency on:**
- How decisions are made.
- Who decides what information is good information?
- How to check decision-making processes.
- Built-in bias and its consequences.
- Security of date storage and management.
- Accuracy of data collection.

Political culture and distribution of power

- The potential for abuse and manipulation.
- Robustness of democratic institutions.
- Distribution of power.
- Future political system.
- Future of international and global alliances.
- Distribution of power and access to communication between private and public actors.

Equality and fairness

- AI driven by data collection, and access to data will drive access to (good quality) AI.
- Application of AI in government and welfare: who should be monitored?
- What safeguards should be implemented for AI application in government administration?
- Increase or decrease in extremism
- Uncertainty if better oversight on issues of debate.
- Protection of individual rights.
- Quality and access to public dialogue and debate.
- Overview of data collection practices.
- Quality of AI systems and application.

JUSTICE & ETHICS

In their discussion on the legal and ethical implications of AI, the experts elaborated on issues of: Legal framework and good data governance; transparency, social and moral responsibility and legal options of enforcement; privacy, self-determination and equality.

Legal framework and good data governance

- Lack of overview.
- Data spread and use will intensify.
- Biased and wrongful profiling.
- GDPR is easily circumvented.
- Anonymity, informed consent and personal security is challenged.

Transparency, Social and moral responsibility and legal options of enforcement

- Holding data platforms and companies accountable for involvement in criminal acts, social division, undermining democracies (elections, debates, tax).
- IP laws and proprietary rights regulation protects companies.

privacy, self-determination and equality

- Changing conditions for and meanings of 'privacy'.
- Cost saving exercises promoted as improvement of e.g. healthcare.
- Increased surveillance.
- Social framework for technology use.
- Discrimination and stigmatization.
- Misuse of data by companies.
- Problems with implementation and execution of GDPR.
- Involuntary voluntariness (opting out of social media use an actual option?).
- Lack of understanding on being good data-citizen/data-citizenship.

ECONOMY

In the expert discussions on AI and related economic implications, the experts elaborated on issues of: expenses for business and private life and conditions of competition, access to data and advancement of research and innovation, and externalities and fair distribution of costs and benefits.

Expenses for business and private life and conditions of competition

- High costs associated with a wide implementation of AI-enabling infrastructures in society.
- Securing good conditions of competition.

Externalities and fair distribution of cost and benefits

- The future will bring a challenge to include qualitative measures and not just settle for the quantitative, when measuring economic growth.
- It is undeniable that productivity in an AI future will increase; however, the salaries might not, and the nature of the needed labour might change.
- A future of haves and super-haves will challenge us to find ways to a fair distribution of labour and wealth.

Access to data and advancement of research and innovation

- Improved access to data for research institutions and companies would stimulate R&I, but how do we share and use data, without forcing unnecessary burdens on individuals?
- There is an imbalance between the public and private sectors' access to data.
- Fear that AI could bring a jaded and subjective spin on R&I and the meaning and quality of data.
- Future AI development is in need of a more multi-disciplinary approach.

SOCIAL IMPLICATIONS

The expert group debating the social implications of AI centred on the themes of the putative futures for labour and the job market in an AI mediated economy, the implications of an AI mediated healthcare system and the implications of an AI mediated educational system.

AI and the future of labour

- Safeguarding labour.
- Creation and meaningfulness of labour.
- The human cost of an AI influenced job-market.

AI and the future of education

- Increasing commercialisation of the education system.
- High focus on monetary value of education.
- A fundamental misunderstanding between successful integration of digital tools and the improvement of education, teaching and learning.
- Breaking down of social structures for managing learning and teaching environments.
- Breaking down of hierarchical structures and failing to prepare students for social and work life.
- Reinforcement of existing inequalities
- Learning how to manoeuvre in an abundance of information.

AI and the future of healthcare

- Interpreting data correctly and achieving the necessary skills to operate AI based machines (e.g. for radiologists) as well as ensuring a high level of education.
- Feeding AI with correct and adequate data.
- Standardisation might exclude huge groups of patients with orphan diseases.
- Coming to terms with the fact that AI will work in some areas but not in others. If you do not have an adequate understanding of a given disease, AI will not do any good
- Ensuring that AI works in tandem with human.
- Not becoming overly focused on funding AI solutions at the expense of low-tech solutions, such as interpersonal communication, to the extent that these provide better/equally preferable results.
- AI has apparently become a semantic strategy for achieving funds. Putting an end to this development is also an ethical question. We need to (collectively and politically) compile a list of 'nice to haves' and 'need to haves' and manifest it in law.
- AI is not better than the data which is fed into the AI: Correct and adequate data is the pre-condition for successful utilisation of AI in diagnosis.

About the workshop

AI 360 COPENHAGEN was established in the belief that it is indeed possible to influence how technology and related societal phenomena should develop. AI entails big hope, big hype and big risk, and an obvious place to begin is to form an overview of where the true hotspots are and what we can do about them. This is where the 360 degrees perspective comes in: In all modesty we set out to reach an inclusive, systematic and comprehensive overview. To do so, we invited experts in the areas of AI related rights and ethics, economy, legal/judicial matters, political significance and societal implications. Their task was, in a structured process, to put their insights and ideas into words, for others, further down the line, to be able to put these words into action.

The name '360' refers to the aim of an all-encompassing approach to the topic of AI; an ambition to achieve "a 360 degrees overview". The methodological inspiration for this task came from the DESSI project (http://securitydecisions.org/) a 2013 EU-funded project developing a process and decision support system aimed at end users of security investments. The DESSI method features a decision-making tool allowing for comprehensive assessment of the potential and consequences of various security dispositions. By analysing these in a systematic and structured manner, the method contributes to a much clearer overview and allows for a transparent and participatory decision-making accounting for context and societal multi-dimensionality in choosing the right investments. The 360 tool invented for and featured in the 360 COPENHAGEN workshop was lifted from the DESSI tool, but much redesigned and adapted to the specificities of AI as a technological and societal phenomenon.

The mission of AI 360 COPENHAGEN is to create an overview of hotspots and possible actions. The 360 COPENHAGEN workshop was the starting point of this endeavour but does not stand alone. The results of the workshop will inform and feed into a European citizen consultation in the summer and autumn of 2019, in which citizens all over Europe will provide their assessment of how society should deal with the AI future. The material they will be discussing are coming straight out of the 360 COPENHAGEN workshop.

About the AI 360 method

At the core of the AI 360 method is the 360tool, which was developed exclusively for the workshop. However, the tool certainly has potential to cover similar evaluations of technological impact and development in other areas, where thorough technological assessment is on the agenda.

The 360 tool is the materialization of a walk-through method for decision support. The tool enables a versatile assessment process of complex societal dimensions of prospective future technological development; here of AI development and implementation in industry, defence, civil society and administration as well as other spheres of political and social life. The 360 tool makes it possible for the user to visualize, address and evaluate different technological futures in a structured yet creative manner. The exercise consisted of looking into several dimensions: impact on rights and ethics, legal frameworks, social implications, political significance, and economy. For each of these dimensions a set of criteria were pre-selected in order to focus the discussion (see workshop report for further details). The aim of this framework is exactly to achieve the beforementioned 360 degrees overview, to avoid the limited scope of much AI popular discourse: When visualizing AI tech futures and advantages of potential smart solutions, the advantages of automation and outsourcing of societal and repetitive tasks and day-to-day decision making often overshadows other important aspects of social and political life, such as e.g. individual rights, autonomy, problems of surveillance, privacy, transparency and equality, freedom from discrimination as well as other significant social, political and economic implications. Debating and assessing a given technology according to the 360 degrees method may certainly affect the thinking and framing of various societal and practical domains such as transport, public space, health care etc. in terms of the way the future of such institutions is imagined.

Authors of this newsletter

Aske Palsberg , Lise Bitsch and Sita Ramchandra Kotnis, The Danish Board of Technology Foundation

Contributing authors:

Nicklas Bang Badum, Marie Louise Jørgensen and Lars Klüver, The Danish Board of Technology Foundation.