Workshop on Ethical aspect of the use of AI in decision making

Arranged by the AI, Ethics and Philosophy Research Group (AEP)
in collaboration with Center for Sustainable ICT (CESICT) and Programme for Applied Ethics (PAE)

Place/time: Scandic Nidelven hotel 11 October, 2023 09.30 – 16.30

09.30: Coffee/tea, mingling

10.00 – 10.30: Introduction (May Thorseth & John Krogstie (AEP and CESICT))

10.30 – 11.15: Ethical issues with BIAS and Mitigating diversity of AI in the Labor Market (Roger Søraa, NTNU)

11.15 – 11.30: Coffee/tea

11.30 – 12.30: Artificial intelligence, moral emotions and ethical decision making (Keynote Sabine Roeser, TU Delft)

12.30– 13.30 Lunch

13.30 – 14.15: AI, Democracy and Political Epistemology: AI as a Threat to Epistemic- and Political Agency
(Heine Alexander Holmen, NTNU)

14.15 – 15.00 NorGLM: Norwegian generative language models and an investigation on toxicity concerns (Lemei Zhang/Peng Liu NorwAI/NTNU)

15.00 – 15.30 Using AI for research. New ethical challenges? (John Krogstie)

15. 30 – 15.45 Coffee/tea

15.45 – 16.30 The Norwegian Government invests billions in artificial intelligence – preparing for a proposal from AEP

17.00 Dinner

Abstracts:

Sabine Roeser, TU Delft; s.roeser@tudelft.nl

Artificial intelligence, moral emotions and ethical decision making

The recent fast advances of artificial intelligence (AI) have become a major issue in societal debates, technology development and ethics research. While AI comes with many promises, it also introduces unprecedented ethical problems. Leading technology developers have warned about the existential risks that AI could pose in the future, proposing a moratorium on its development. However, critics have argued that this warning may further fuel the hype around AI and distract from more mundane problems with AI that need much more urgent attention. These issues concern (next to e.g. environmental problems, due to huge energy consumption in the design and use of AI) the impact AI already has on human decision making, by introducing and enforcing biases and unfairness as well as lack of transparency. In this contribution I will argue that moral emotions such as sympathy, care and feelings of responsibility can help us to draw attention to these risks of AI systems and lead to insights into better designs. Furthermore, I will argue that the lack of emotions and ethical stance of AIs contributes to their problematic features. I will discuss some recent examples, such as the Dutch tax benefits scandal that led to the fall of the Dutch government in 2021. The Dutch tax authorities used an AI system that was based on racist assumptions and a reversal of the principle 'innocent until proven guilty'. This led to unsurmountable Kafkaesque bureaucratic hurdles and personal tragedies for countless families. Promised compensation for the repercussions are equally lost in bureaucratic conundrums. This case serves as an illustration and warning of problems that AI systems can create. In order to address these challenges, a humane, caring approach is needed and should form the basis of any employment and development of AI. This also means that we need strict requirements for governments and businesses using and developing AI.

Roger Søraa, NTNU

Ethical issues with BIAS and Mitigating diversity of AI in the Labor Market

This talk will address issues of bias for AI technologies in the labor market, exploring its implications on diversity and inclusivity. Practical strategies for mitigating these biases will be explored, emphasizing the creation of fair and equitable AI tools. I will focus on the importance of diversity in AI development processes, for implementing solutions that promote equality and prevent discriminatory practices within the rapidly evolving labor market landscape.

Heine Alexander Holmen, NTNU

AI, Democracy and Political Epistemology: AI as a Threat to Epistemic- and Political Agency

ABSTRACT: Most theories of democracy assume that citizens of a well-functioning democracy must possess certain epistemic capacities that in turn enable its citizens to form and acquire the politically relevant knowledge-cum-information required for sound democratic processes

– viz. either for voting, for engaging more directly in democratic deliberation, or for contributing to other sorts of participatory political practices. Paradoxically as it may sound, however, AI may pose a threat to these epistemic capacities of the citizens in virtue of being a strong epistemic tool for information processing and knowledge formation. The reason is that the widespread and unchecked usage of AI, in all its splendour and glory, may risk undermining the epistemic agency of citizens by virtue of discouraging their trust in one's own epistemic capacities. Lack of epistemic trust in one's epistemic capacities may in turn lead to less exercise of those capacities among citizens with the result that the community over time risks undermining the epistemic agency of its citizens. Consequently, AI could undermine the very epistemic capacities required for political agency and democratic participation. In the talk, I explore these threats and argue that there are strong reasons for taking them seriously in a democracy. However, I also argue that there are means for addressing them – especially through education and educational policies.

Lemei Zang/Peng Liu, NTNU:
NorGLM: Norwegian generative language models and an investigation on toxicity concerns

Norwegian, spoken by just 5 million people worldwide, is underrepresented in NLP research. Previous work on Norwegian has been hampered by a lack of annotated datasets, a scarcity of language resources, and a lack of resource standardization. In this presentation, we will introduce NorGLMs, a collection of foundational Norwegian generative language models trained on a newly created 196 GB Norwegian corpus. We will report the performance of our models and compare with another Norwegian language model on a set of standard benchmarks. Finally, we expose some of the biases and toxicity encoded in our models, sparking further discussions, and raising concerns in this rapidly evolving field.

John Krogstie, NTNU:
Using AI for research. New ethical challenges?

Accelerating the productivity of research could be the most economically and socially valuable of all the uses of artificial intelligence (AI). We have some star-examples of the use of AI in research, such as Alphafold predicting protein-folding. A lot of the use of AI in research accentuate issues related to the use of data and openness needed about the research workflow, research results and research data, but there are also new issues arising. While AI is penetrating all domains and stages of science, its full potential is far from realized. Policy makers and actors across research systems can do much to accelerate and deepen the uptake of AI in science, magnifying its positive contributions to research, balanced by measures to take new and already present ethical issues into account.

Participants:

Gitte Koksvik	gitte.koksvik@ntnu.no	NTNU, Soc. Anthropology
Anders Nes	anders.nes@ntnu.no	NTNU, IFR
Heine A. Holmen	heine.a.holmen@ntnu.no	NTNU, IFR
Espen Stabell	eds@hvl.no	Hvl/NTNU Business Sch.
Hitesh	hiteshh@ntnu.no	NTNU Econ./ Managem.
Roger Søraa	roger.soraa@ntnu.no	NTNU, KULT
Aurora Hoel	aurora.hoel@ntnu.no	NTNU Art/Media
Silvia Ecclesia	silvia.ecclesia@ntnu.no	NTNU, KULT
John Krogstie	John.krogstie@ntnu.no	NTNU, Comp. Science
Jonathan Knowles	jonathan.knowles@ntnu.no	NTNU, IFR
May Thorseth	may.thorseth@ntnu.no	NTNU, IFR
Lemei	lemei.zhang@ntnu.no	NTNU, Comp. Science
Peng Liu	peng.liu@ntnu.no	NTNU, Comp. Science
Ronny Selbæk Myhre	ronny.s.myhre@ntnu.no	NTNU, IFR
Nora Johanne Klungseth	nora.klungseth@ntnu.no	Mech./Ind. Engineering
Tobias Brox Bordal	Tobiasbb@ntnu.no	NTNU, IFR
Ahmed Idries	ahmed.y.m.idries@ntnu.no	NTNU, Comp. Science
Jingyue Li	jingyue.li@ntnu.no	NTNU, Comp. Science
Eric Monteiro	eric.monteiro@ntnu.no	NTNU, Comp. Science
Pieter Toussaint	Pieter@ntnu.nu	NTNU, Comp. Science
Syed Sajid Hussain	syed.shah@ntnu.no	NTNU, Comp. Science
Ilaria Crivellari	ilaria.crivellari@ntnu.no	NTNU, Comp. Science
Durga Prasad Bavirisetti	durga.bavirisetti@ntnu.no	NTNU, Comp. Science
Dag Svanæs	dags@idi.ntnu.no	NTNU, Comp. Science
Eirik Flogard	eirik.l.flogard@ntnu.no	NTNU, Comp. Science
Adrian Sand	Adriasa@stud.ntnu.no	NTNU, Comp. Science
Alicia Takaoka	alicia.j.w.takaoka@ntnu.no	NTNU, Comp. Science
Mathias Lundteigen Mohus	mathias.l.mohus@ntnu.no	NTNU, Comp. Science
Charlotte Grøder	charlotte.h.groder@ntnu.no	NTNU, Comp. Science

Johanna Johansen	johanna.johansen@ntnu.no	NTNU, Comp. Science
dina borge	dinamb@ntnu.no	NTNU, Soc. Anthropology
Simon A. Berger	simon.berger@ntnu.no	NTNU, Edu./Lifel. Learn.
Sven Herman Holmsen	svenhho@stud.ntnu.no	NTNU, Comp. Science
Kirsti E Berntsen	kirsti.berntsen@ntnu.no	NTNU, Comp. Science
Jussi Haukioja	jussi.haukioja@ntnu.no	NTNU, IFR
Celina Treullier	celina.treullier@loria.fr	NTNU, Comp. Science