2025 11 06 Doctoral student day, Mälardalen University

AI ETHICS

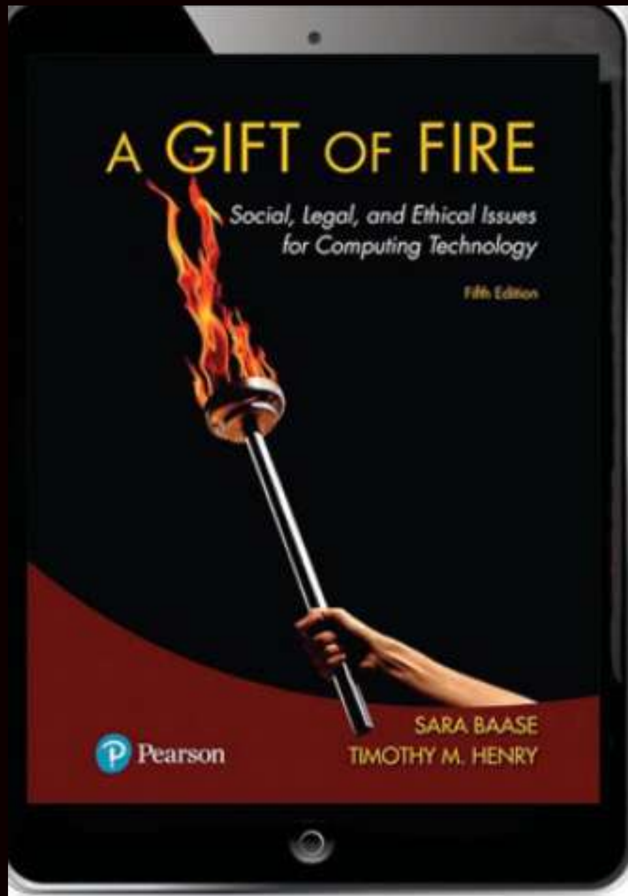
Multiperspective Ethics in the Age of AI

Gordana Dodig Crnkovic

Chalmers University of Technology | Unuversity of Gothenburg & Mälardalen University, Västerås,

gordana.dodig-crnkovic@chalmers.se
gordana.dodig-crnkovic@mdu.se

https://www.gordana.se/



Affiliations

School of Innovation, Design and Engineering

Division of Computer Science and Software Engineering

Research groups:
Artificial Intelligence and Intelligent Systems
Ubiquitous Computing

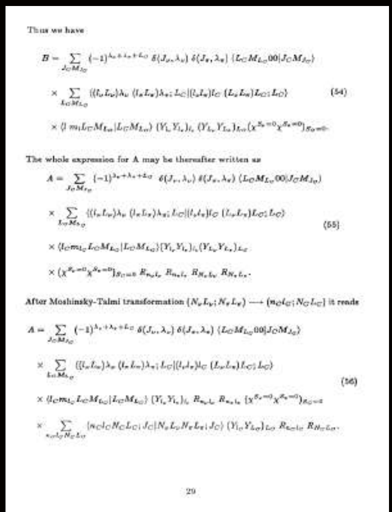
AI & Society-AI@MDU Research Fellow

Department of Computer Science and Engineering

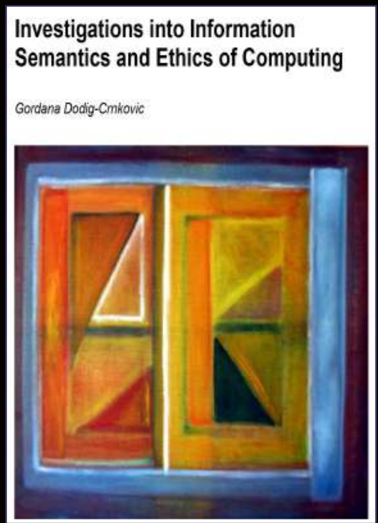
Computer Science and Software Engineering Division

Research groups:
Interaction Design and Software Engineering
Critical Robotics

Background



PhD in Physics, 1988
On Alpha-decay, Department of Physics, University of Zagreb



PhD in Computing, 2006
Mälardalen University

Current: Morphological Computing and Cognition

AI Ethics, Digital Ethics, Digital Humanism

Multiperspective view of ethics

- Member and Chair of MDU's Ethics Committee
- Member of Chalmer's AI Ethics Committee
- Researcher in two complementary research fields:
  computer science and philosophy, where ethics is a central concern
- Teching cethics ourses ***Professional Ethics in Science and Engineering*** (2003 – 2017) (PhD c& Masters course) at MDU , ***Research Ethics & Sustainable Development***, PhD Course at Chalmers University of Technology (2014-2017) , ***Professional Ethics Digital Ethics and the Connected World*** as Guest Professor at TUW (2023 – 2024) and numerous ethics lectures in other courses like Research Methodology course, Emerging Trends in Interaction design, Software Engineering, Cognitive Science Methods, etc.)
- Working in projects to improve AI introduction in pedagogical context, currently involved in a research project between MDU, LiU, and KTH on integrating AI tools into teaching

Multiperspective view of ethics

- Board member of Informatics Europe, leading Ethics group, organizing workshops on ethics
- Associate Editor of the journal Philosophies and board member of several journals, where we are developing publishing policies in response to the rapid rise of AI
- Board member of the International Academy for Information Studies, focusing both on the mechanisms behind AI and its broader impact on society
- External member of The Karel Capek Center for Values in Science and Technology, a research multidisciplinary center established by the Czech Academy of Science and the Charles University, Prague.
- Research fellow of AI & Society-AI@MDU
- Citizen and private individual — it's impossible not to notice the changes in everyday life brought about by AI

Terminological Clarification: Ethics and Morality

The terms ethics and morality are often used interchangeably - indeed, they usually can mean the same thing, and in everyday conversation there isn't a problem with switching between one and the other.

However, there is a distinction between them in philosophy!

Ethics and Morality, Etymology

Morality and ethics have the same roots, mores which means manner and customs from the Latin and etos which means custom and habits from the Greek. (Robert Louden, Morality and Moral Theory)

Strictly speaking, morality is used to refer to what we would call moral conduct while ethics is used to refer to the formal study of moral conduct.

Ethics is also often called moral philosophy.

Ethics and Morality, in Short

MORALITY - PRACTICE: first-order set of beliefs and practices about how to live a good life.

ETHICS - THEORY: a second-order, conscious reflection on the adequacy of our moral beliefs.

Ethicist Prof. Ibo van de Poel from TU Delft in the Netherlands suggests that the students need to develop the following "moral competences":

Moral sensibility

Moral analysis skills

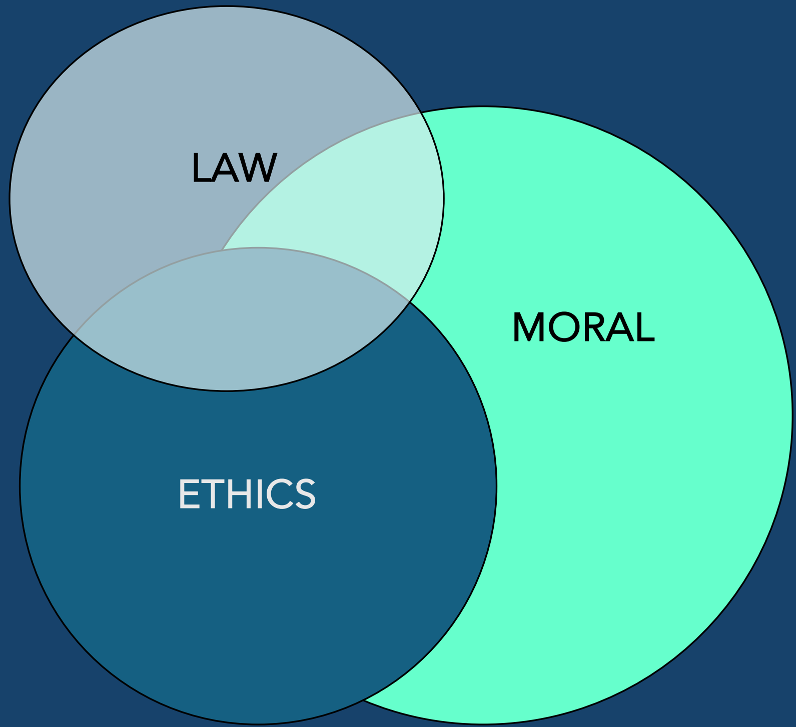
Moral creativity

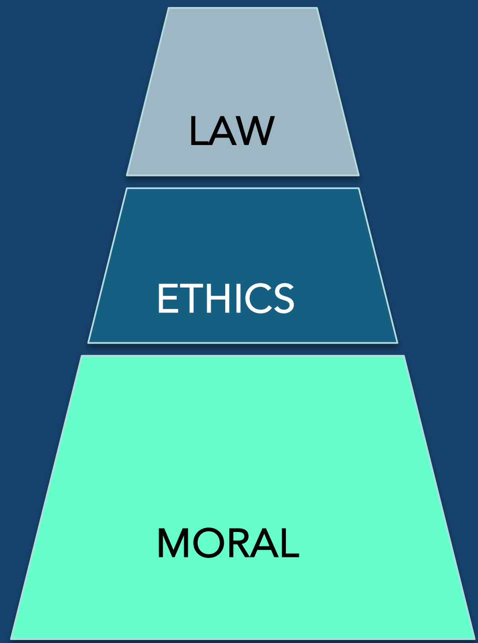
Moral judgment skills

Moral decision-making skills

Moral argumentation skills

Societal Normative Systems



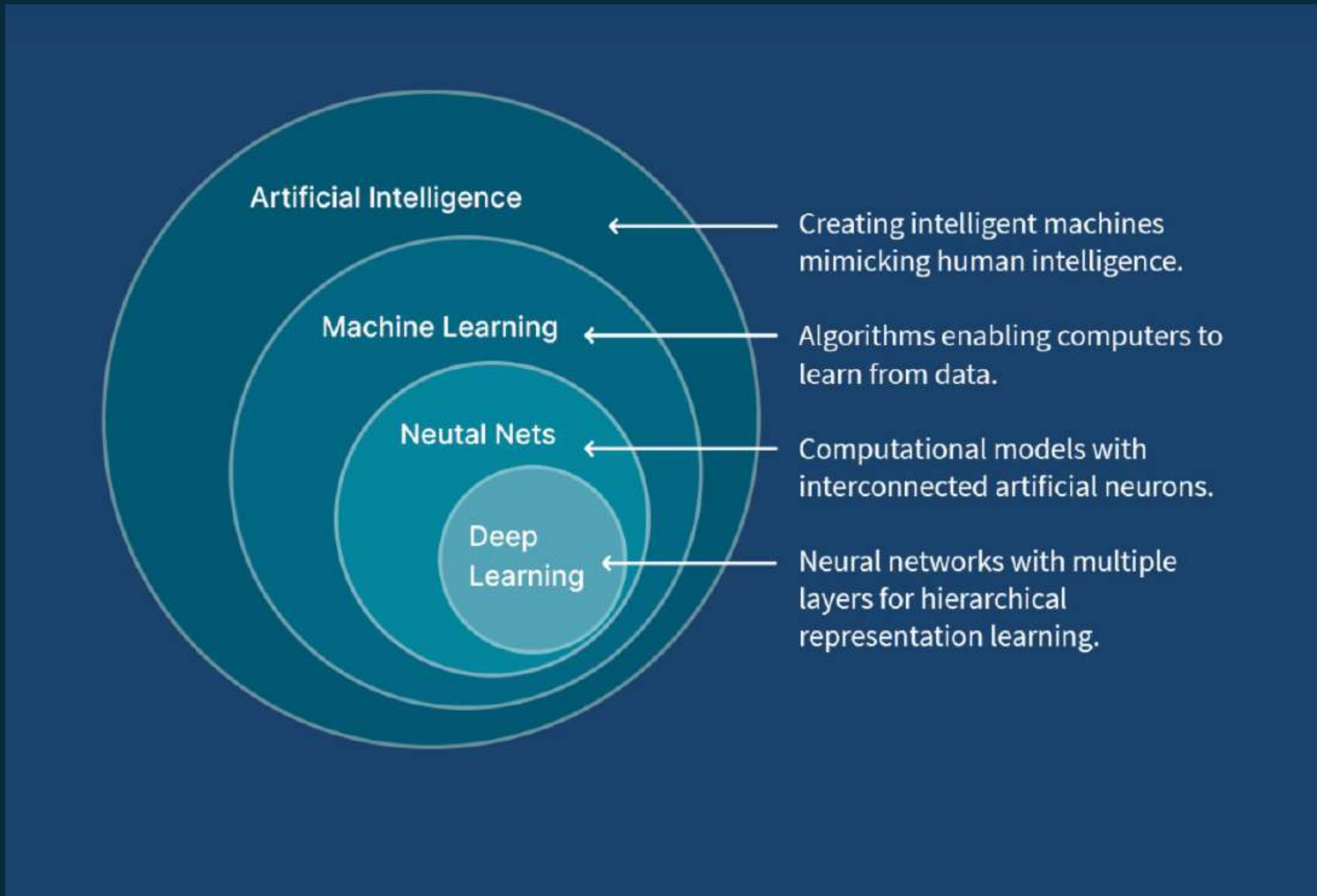


Ethics as Continuum - An Ongoing Conversation

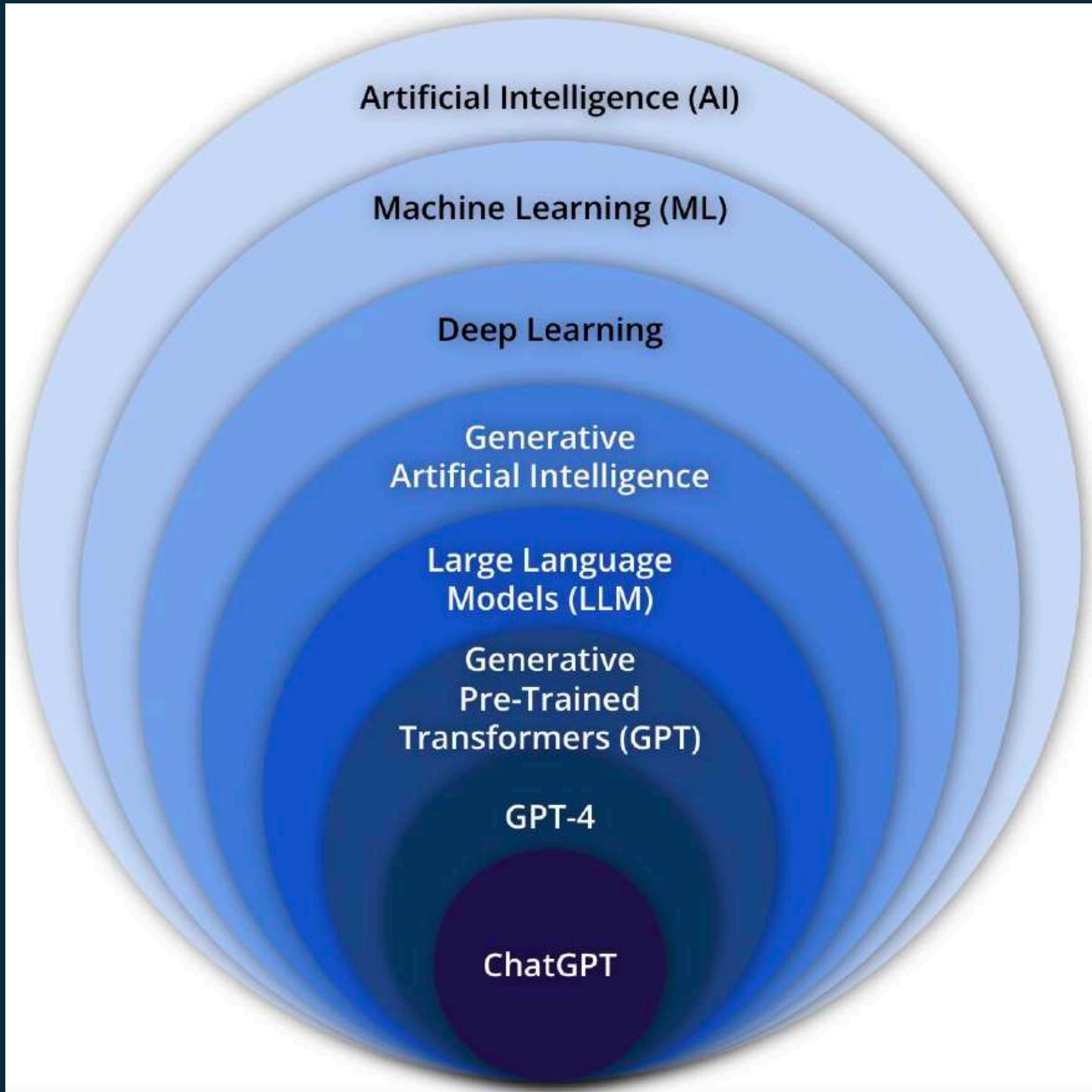
- World changes constantly, and we have to interpret/construe it over and over again.
- We come back to ideas again and again, finding new meaning in them.
- Professional discussions of ethical issues in journals.

See http://www.utm.edu/research/iep/e/ethics.htm Ethics

WHAT IS AI?


Artificial Intelligence
Machine Learning
Neutal Nets
Deep Learning
Creating intelligent machines mimicking human intelligence.
Algorithms enabling computers to learn from data.
Computational models with interconnected artificial neurons.
Neural networks with multiple layers for hierarchical representation learning.

WHAT IS AI?



Artificial Intelligence (AI)
Machine Learning (ML)
Deep Learning
Generative Artificial Intelligence
Large Language Models (LLM)
Generative Pre-Trained Transformers (GPT)
GPT-4
ChatGPT


Even though AI is much bigger field, we will focus on LLMs

Launched in November 30, 2022, ChatGPT and similar LLMs are here to stay

The Early Foundations GenAI (2016–2021) by Open AI

The ChatGPT Launch (2022) On November 30, 2022, when OpenAI launched it as a free research preview.

Google Bard came in March 2023 and

Anthropic Claude also launched in 2023.

Explosive growth since then

Read more:
https://tattvammedia.com/blog/history-of-chatgpt/

STATE OF THE ART, AI APPLICATIONS IN RESEARCH AND EDUCATION

Research Council to Investigate How AI Can Be Used in Research (October 1, 2025)

The government has tasked the Swedish Research Council (Vetenskapsrådet) with promoting the use of artificial intelligence (AI) in research. The assignment involves mapping needs, opportunities, and obstacles, as well as analyzing how AI can be utilized as a scientific tool to strengthen the conditions for world-leading research in Sweden.

In the 2024 Research and Innovation Bill, the government highlights AI as a technology with the potential to revolutionize scientific methods. The EU has also recently emphasized the importance of a common strategy for AI in research.

"We see that AI tools are already being used in research today and will play an increasingly significant role across all fields. Our task now is to create a comprehensive overview of the needs and to help ensure that AI is used in an effective and responsible way," says Viktoria Mattsson, Head of the Collaboration Department at the Swedish Research Council.

As part of the assignment, the Swedish Research Council will identify the need for support and guidance for researchers and carry out an international analysis of external developments. The goal is to enable a transparent, ethical, and efficient use of AI throughout the research process—an approach aligned with principles of scientific excellence, openness, and research integrity, and one that does not increase administrative burden or create unnecessary regulatory complexity.

Research Council to Investigate How AI Can Be Used in Research (October 1, 2025)

Dialogue with Universities, Funders, and Authorities

The work will be carried out in close dialogue with universities, other research funders, and relevant authorities. Vinnova will contribute with expertise and experience, especially related to its mission concerning strategic technologies.

The Swedish Research Council is already running other initiatives where AI plays a central role. These include work on research e-infrastructure and the "Excellence Clusters for Groundbreaking Technology" program. Therefore, it is important that the new AI initiative is coordinated with these efforts.

"What is happening within these initiatives can provide valuable insights into how AI is already being used in research and thus form an important foundation for the new assignment," says Viktoria Mattsson.

The work will take place from 2026 to 2028 and will be reported annually to the Government Offices.

Assignment to the Swedish Research Council to nationally coordinate efforts to promote the use of artificial intelligence in research

https://www.vr.se/english/just-now/news/news-archive/2025-10-08-the-swedish-research-council-to-explore-how-ai-can-be-used-in-research.html

Research Council Guidelines for the use of AI tools

Developments in artificial intelligence (AI) are rapid, not least due to research in the subject. Here you can read about the Swedish Research Council's guidelines for using AI tools in applications for research funding and in the assessment of applications.

AI in applications

There is no prohibition against you as applicant using generative AI or other tools (digital or of another type) when you draw up your application. At present, you do not need to state whether you have used AI.

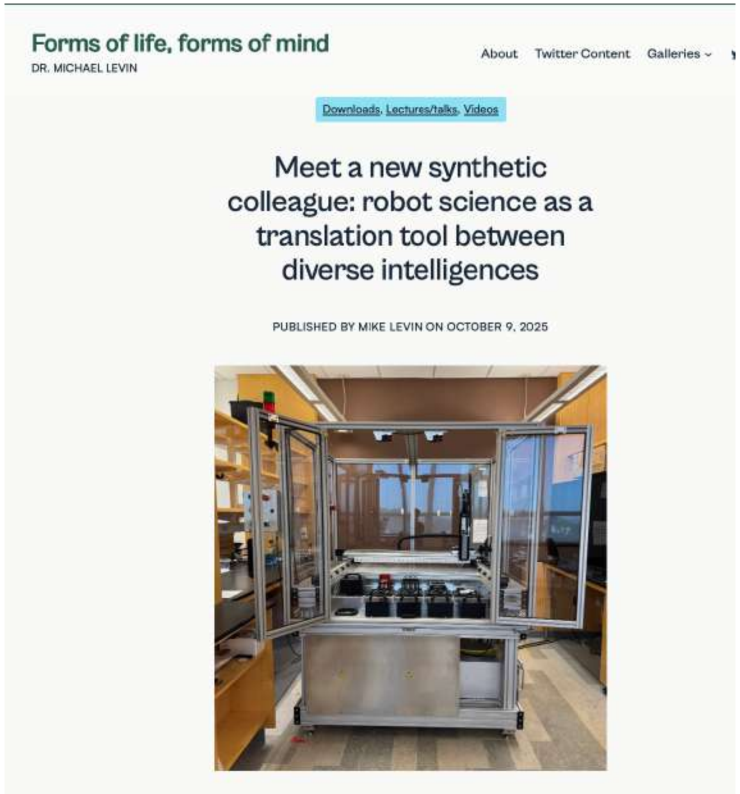
On the other hand, you need to be aware of the following:

As the applicant, you are responsible for ensuring the contents of the application – including everything stated in the project plan – are correct and implementable as described. You will confirm this in conjunction with registering the application.

As the applicant, you must comply with good research practice during the application procedure, which means that plagiarism, falsification, or fabrication of content in the application must not occur.

https://www.vr.se/english/applying-for-funding/applying-for-a-grant/guidelines-for-the-use-of-ai-tools.html

New Robotic Research Assistants



Open AI winning programming competition

OpenAI Just Beat the World's Best Programmers at Their Own Game

The world's top college programmers just got schooled by AI. OpenAI's reasoning models achieved a perfect 12/12 score at the ICPC World Finals, the most prestigious programming competition in the world... *and outperforming every human team.*

To put this in perspective, the best human team solved 11 out of 12 problems. And

https://www.theneurondaily.com/p/ai-s-huge-competitive-coding-win September 18, 2025

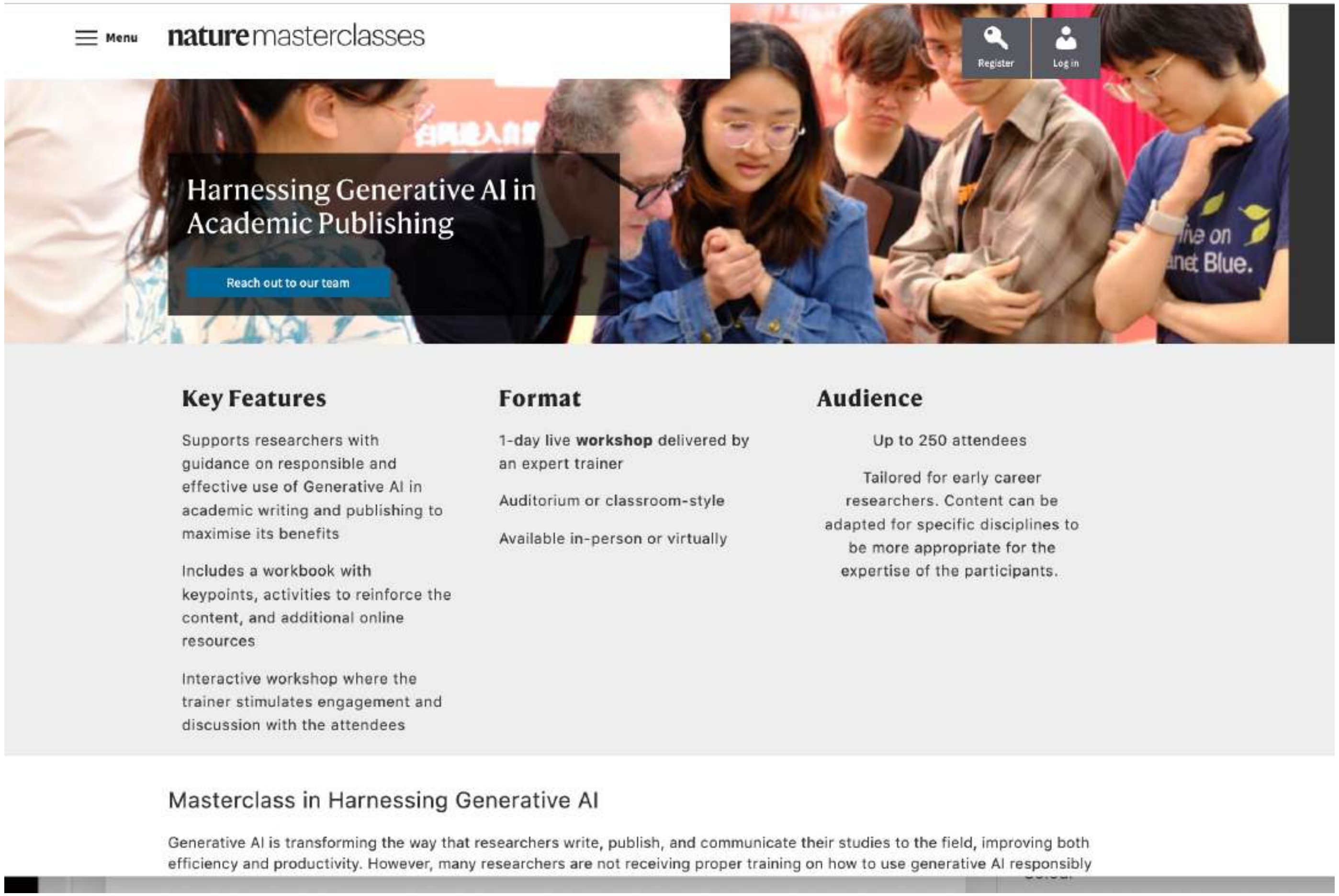
AI Models Analyze Language As Well As a Human Expert

A recent study showed that OpenAI's "o1" language model could analyze complex grammar, recursion, and ambiguity at a graduate-linguist level. This challenges claims that AI lacks true linguistic understanding and suggests large models may be closing in on abilities once thought uniquely human.

“The new results show a steady “chipping away” at properties that had been regarded as the exclusive domain of human language”, Beguš said. “It appears that we're less unique than we previously thought we were.”

https://www.quantamagazine.org/in-a-first-ai-models-analyze-language-as-well-as-a-human-expert-20251031

How to learn using AI the best way


Menu **nature**masterclasses

Register Log in

Harnessing Generative AI in Academic Publishing

Reach out to our team

Key Features

Supports researchers with guidance on responsible and effective use of Generative AI in academic writing and publishing to maximise its benefits

Includes a workbook with keypoints, activities to reinforce the content, and additional online resources

Interactive workshop where the trainer stimulates engagement and discussion with the attendees

Format

1-day live **workshop** delivered by an expert trainer

Auditorium or classroom-style

Available in-person or virtually

Audience

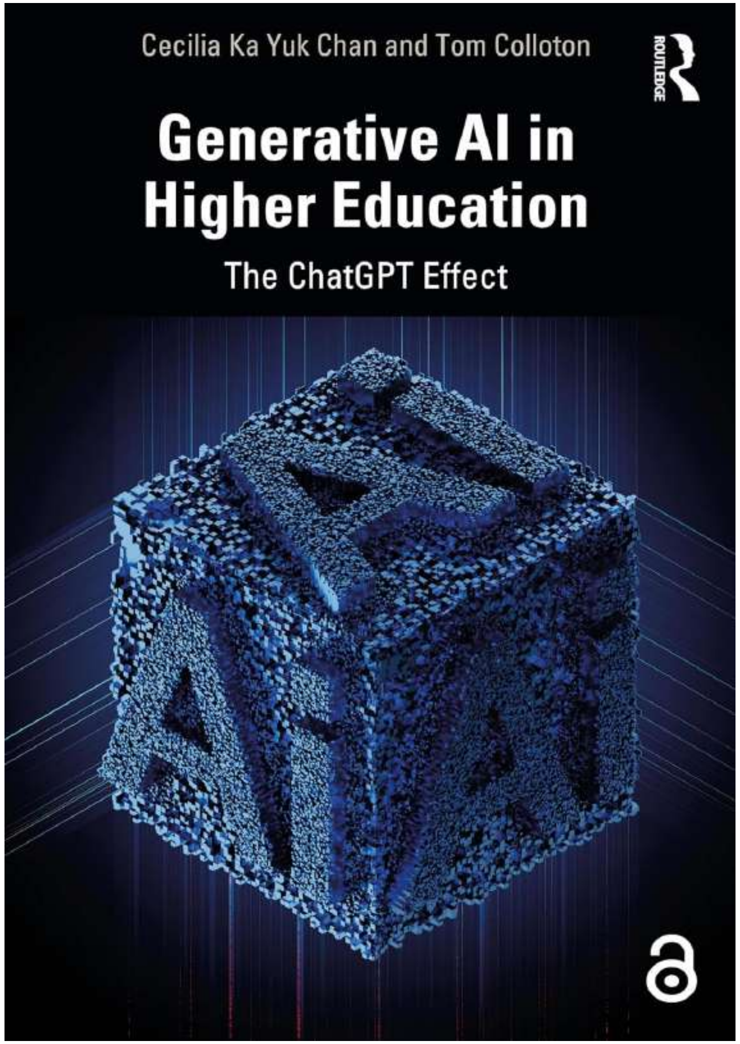
Up to 250 attendees

Tailored for early career researchers. Content can be adapted for specific disciplines to be more appropriate for the expertise of the participants.

Masterclass in Harnessing Generative AI

Generative AI is transforming the way that researchers write, publish, and communicate their studies to the field, improving both efficiency and productivity. However, many researchers are not receiving proper training on how to use generative AI responsibly

https://tinyurl.com/yc6b2twa



Generative AI in Higher Education

The ChatGPT Effect

Open access

Chan and Colloton's book is one of the first to provide a comprehensive examination of the use and impact of ChatGPT and Generative AI (GenAI) in higher education.

https://www.routledge.com/Generative-AI-in-Higher-Education-The-ChatGPT-Effect/Chan-Colloton/p/book/9781032599045

AI ETHICS

Multiperspective Ethics in the Age of AI

A Gift of Fire: Social, Legal, and Ethical Issues for Computing Technology

Sara Baase

First Published: 1997 (the most recent the 5th edition)

This book explores ethical and social issues related to computing and digital technology—covering privacy, intellectual property, freedom of speech, computer crime, and broader societal impacts.

Ancient Greek Myth on Prometheus

The Greek myth of Prometheus tells of a Titan who was a champion of humanity.

He defied the chief god, Zeus, by stealing fire and giving it to humans to enable progress and civilization.

As punishment, Zeus had Prometheus chained to a rock where an eagle would eat his liver daily, with the liver regenerating each night.

Eventually, the hero Heracles freed Prometheus from this eternal torture.

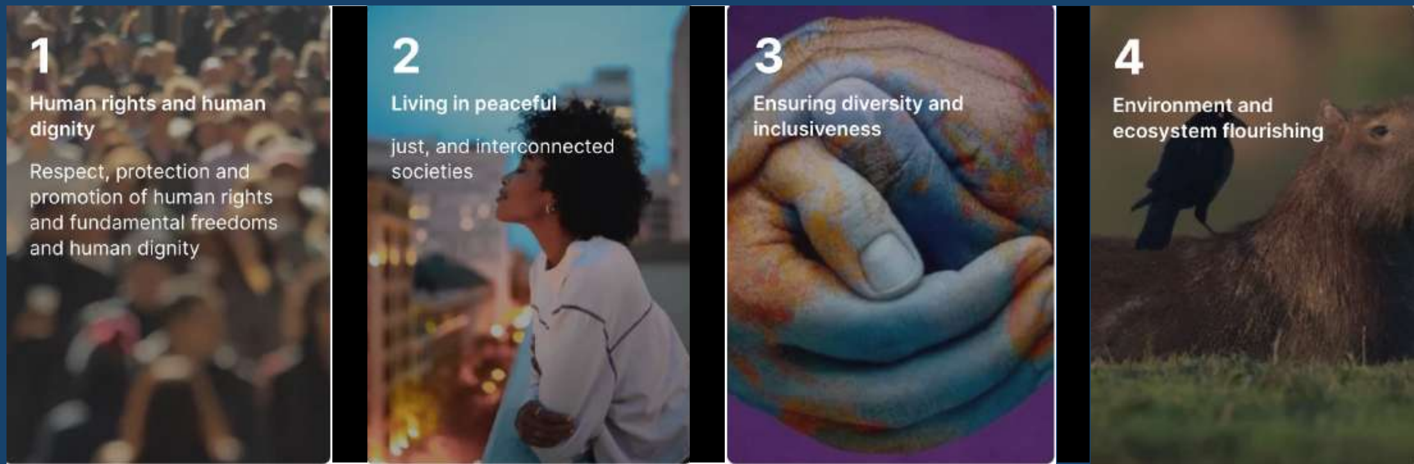




AI IS LIKE A FIRE-JUST MUCH MORE POWERFUL ON THE GLOBAL LEVEL

UNESCO 2022
‘Recommendation on the Ethics of Artificial Intelligence’

UNESCO's first-ever global standard on AI ethics



UNESCO: The United Nations Educational, Scientific and Cultural Organization
https://www.unesco.org/en/artificial-intelligence/recommendation-ethics
https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence

EU "AI Act" (2024)

The world's first AI legislation

AI Act, European Commission. Shaping Europe's digital future

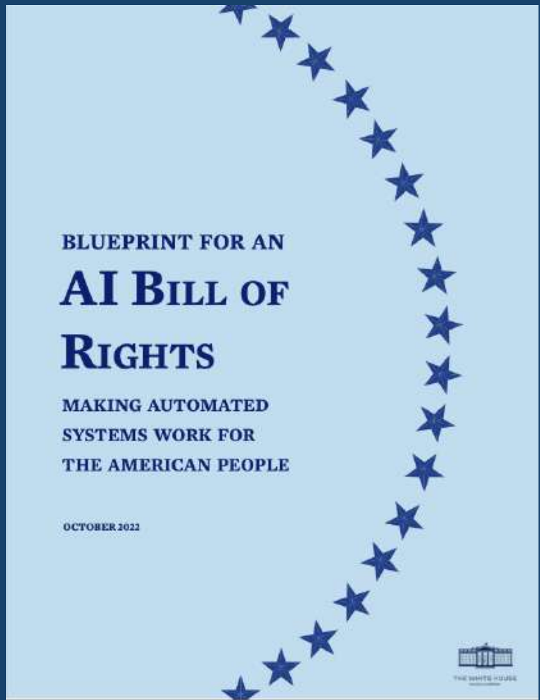
https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

https://artificialintelligenceact.eu/ai-act-explorer/

ISACA

The European Parliament granted final approval of the EU Artificial Intelligence Act on March 13, 2024, by a vote of 523 for passage, 46 against, and 49 abstaining. The Act faces a final step – approval by EU member states – as its provisions gradually take effect.

USA "AI Bill Of Rights" (2022) - retracted



https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf

Currently there is no overarching federal AI regulation, U.S. regulation is patchwork: different states may regulate different aspects, creating complexity for organisations operating across states.

The US AI Bill of Rights outlined principles, including that people have a **right to control how their data is used** and to **not be discriminated against by unfair algorithms**.

It is a white paper, which does not have the force of law. It's primarily aimed at the federal government and could influence **which technologies government agencies acquire**, or help parents, workers, policymakers, and designers **ask tough questions about artificial intelligence systems**.

PROFESSIONAL ORGANISATIONS

IEEE* (Institute of Electrical and Electronics Engineers) on future intelligent autonomous systems

ETHICALLY ALIGNED DESIGN

The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems

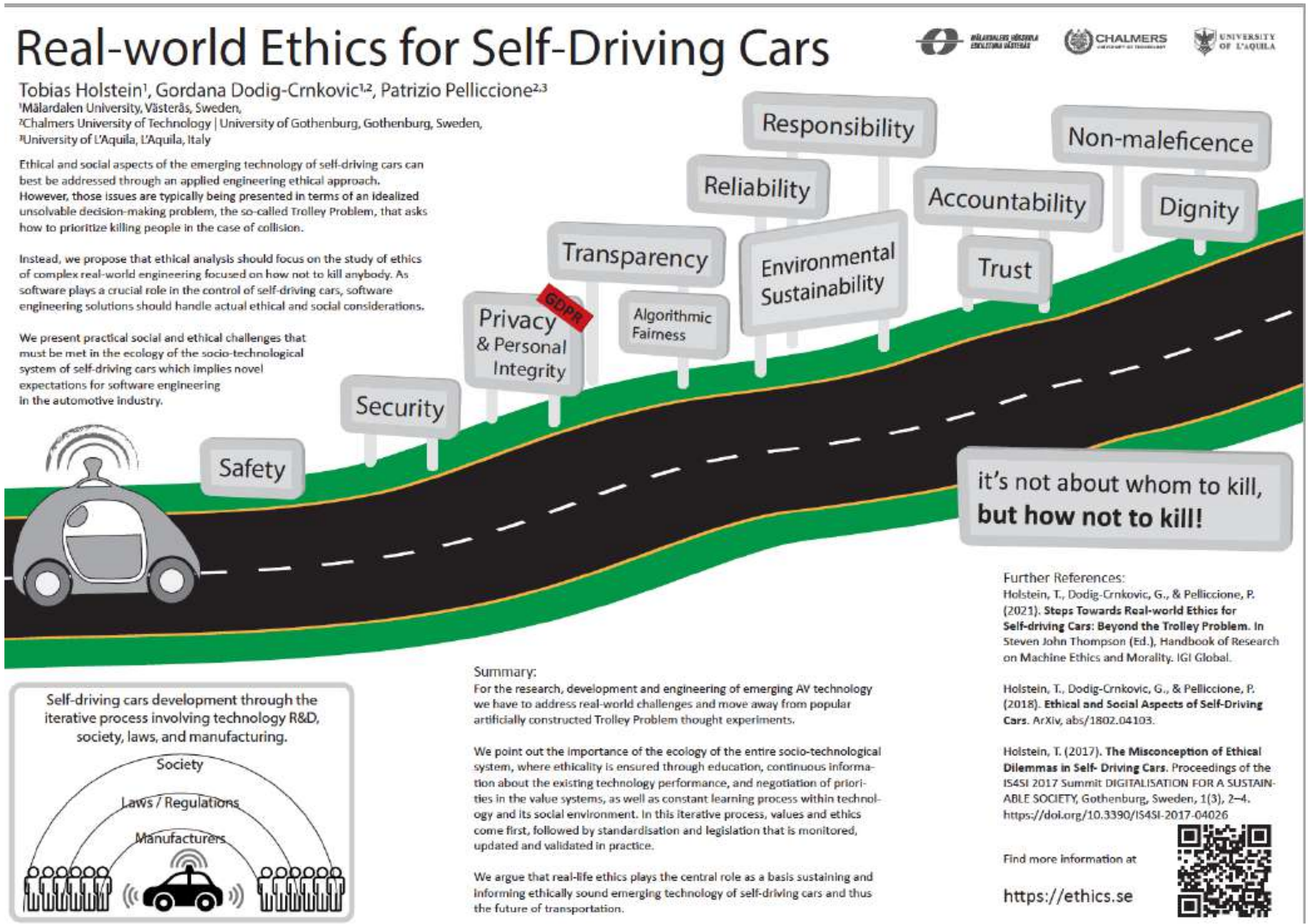
https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf

Prioritizing human well being in the age of artificial intelligence: https://youtu.be/z5yZU8tp9W8 (5:56)

https://techethics.ieee.org/

*IEEE (USA) is the world's largest technical professional organization dedicated to advancing technology for the benefit of humanity.

EXAMPLE: EHICS OF SELF-DRIVING VEHICLES



Presented as poster at ICSE2020
Extended version to appear as a chapter in the
Handbook of Research on Machine Ethics and Morality
| IGI Global 2021

EXAMPLE

Bias, Transparency, and Accountability
in AI-Driven Digital Health Technologies

A Human-Centered Responsibility Approach

Gordana Dodig-Crnkovic

Mälardalen University
Chalmers University of Technology &
University of Gothenburg

https://play.mdu.se/media/t/0_vbp7hfbw

Lecture Roadmap

Main Categories – AI in Health

DIAGNOSTIC TOOLS
Radiology image interpretation (e.g., breast cancer detection)
Dermatology apps for skin lesion analysis
Pathology slide analysis (digital pathology)
Tuberculosis chest X-ray screening in LMICs

THERAPEUTIC SUPPORT & CLINICAL DECISION TOOLS
AI-driven treatment recommendation systems (oncology, cardiology)
Drug dosage optimization (e.g., insulin pumps)
Personalized medicine using genetic profiles

SURGICAL ASSISTANCE
Robotic surgery with AI-enhanced precision
Real-time intraoperative decision support

POPULATION HEALTH & HEALTH SYSTEMS MANAGEMENT
Predictive analytics for hospital readmission risk
Resource allocation & supply chain optimization
Workforce planning in health systems

EPIDEMIOLOGY & PUBLIC HEALTH SURVEILLANCE
Outbreak detection & prediction (e.g., HealthMap)
COVID-19 contact-tracing & triage tools
Environmental health monitoring (air/water quality)

BIOMEDICAL & CLINICAL RESEARCH
Drug discovery & repurposing (AI-assisted vaccine design)
Genomic analysis for disease understanding
Protein structure prediction (AlphaFold)

Core Concepts in AI-Driven Digital Health Technologies Ethics

BIAS IN AI FOR HEALTH

Originates from non-representative, low-quality, or biased datasets

Can lead to misdiagnosis, inequity, and exclusion of groups

Mitigation: representative datasets, bias audits, stakeholder inclusion

TRANSPARENCY IN AI FOR HEALTH

Technical transparency: disclose data sources, model design, and functioning

Procedural transparency: governance processes, documentation, limitations

Explainability at appropriate level for audience (clinician, patient, regulator)

Enables trust, public consultation, and informed debate

PRINCIPLE OF HUMAN-CENTERED RESPONSIBILITY

Final accountability lies with humans, not algorithms

Continuous oversight and bias detection

Decisions must be comprehensible to patients

Legal and institutional enforcement of responsibility

ACCOUNTABILITY IN AI FOR HEALTH

Humans retain ultimate responsibility for outcomes

Clear task definitions, validated scope, and user training

human oversight at design, deployment, and monitoring

Mechanisms for redress when harm occurs

Technologies Promises and Perils

Technology	Promise	Peril
Fire	Warmth, tools, food	Destruction, war
Printing Press	Knowledge for all	Propaganda, misinformation
Industrialization	Productivity, wealth	Inequality, exploitation
Internet	Global access, connection	Surveillance, radicalization
AI	Automation, insight, creativity	Bias, manipulation, power imbalance

Throughout history, powerful technologies have always come with a shadow side. AI is not unique in this, but it is faster, more opaque, and more socially embedded than anything we have seen before.

MANIFESTO OF TODAY - Written by GPT4

The Manifesto for a Humane Future in the Age of Acceleration

1. We Declare Our Right to Reality

We live in an era of manufactured truths, deepfakes, and algorithmic distortion.
We commit to preserving and promoting truth as a public good.
We demand transparency from institutions, platforms, and technologies that shape what we see, believe, and share.

2. We Reclaim Our Time from the Machines

Attention is the currency of modern exploitation.

We will not surrender our minds to infinite scrolls, engineered outrage, and distraction.
We choose intentionality over compulsivity. Rest is resistance.

3. We Center Humanity in All Systems

Artificial intelligence, automation, and data-driven decisions must serve the dignity of all people.
We reject systems that amplify bias, erase cultures, or profit from inequity.
We believe ethical technology is not optional—it is foundational.

The Manifesto for a Humane Future in the Age of Acceleration

4. We Are Stewards, Not Owners, of the Earth

Climate collapse is no longer a distant warning.
We will live, build, and consume as if future generations matter—because they do.
We hold accountable those who treat nature as disposable and extraction as entitlement.

5. We Embrace Collective Intelligence

No one person, algorithm, or nation holds all the answers.
We believe in the wisdom of diverse communities and cross-cultural cooperation.
We choose dialogue over division. We are not adversaries—we are collaborators.

6. We Defend the Margins

Progress that excludes is regression in disguise.
We will amplify voices long silenced and center justice in all pursuits.
We believe equity is not charity—it is overdue correction.

The Manifesto for a Humane Future in the Age of Acceleration

7. We Are Not Passive in Our Future

Silence is complicity. Cynicism is surrender.
We will not outsource agency to systems that profit from our disengagement.
We choose action over apathy, even when the outcome is uncertain.

This is a Manifesto for Now.

A reminder that technology is a tool, not a destiny.

That truth still matters.

That we still matter—and so do our choices.

We write this not as prophets, but as participants.

Not as saviors, but as citizens.

We are already shaping the future.

Let us shape it wisely, together, and humanely.

Would you like a version of this manifesto tailored to a specific domain—education, AI, environment, or governance?

The question is: where does this all come from?

From the data it was trained on,

and reinforcement learning from human feedback.

Good to remember and not mystify GenAI.

They collect huge amounts of information, compressed, and when prompted, they are unpacking and relating it for us. They have much more information than any individual human can have.

But information must be interpreted and valued to become knowledge.

Ethics is not a checklist. It is beyong compliance. It is supporting sustainable democratic society

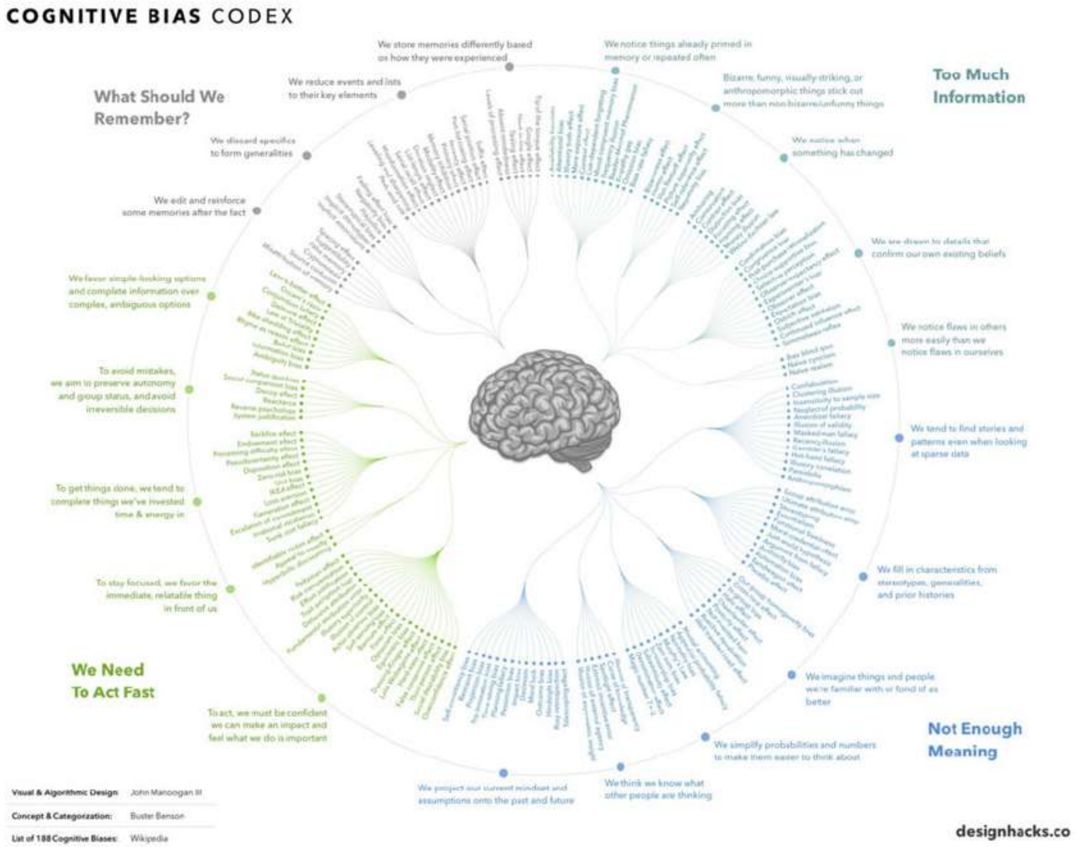
Ethics is not a checklist.
It is an ongoing conversation, that must include more than engineers and ethicists.
It must include all stakeholders,people that are affected, and those who are affecting the phenomenon it is studying.
Ethics is not the “brake” on science and technology.
It is the steering.

Ethics is not about slowing progress or stopping innovation.
It is about guiding it in the right direction.

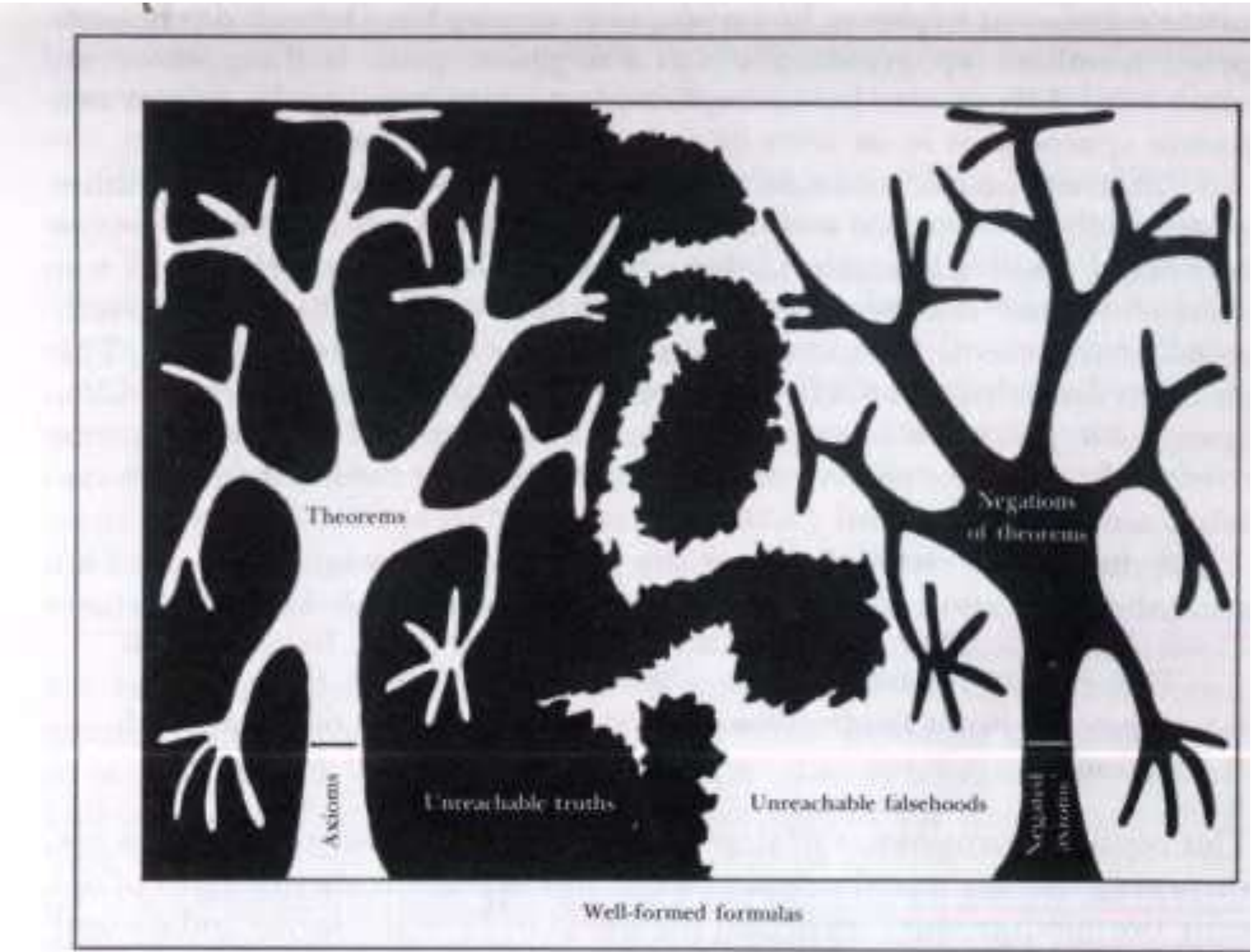
HOW IS ETHICS DONE: STEPS IN ETHICAL ANALYSIS

1. WHAT IS THE CASE?
   Taking into account that our knowledge is limited/imperfect.
   How can we interpret it in the context?
2. STAKEHOLDERS & OBLIGATIONS
3. RELATED NORMATIVE SYSTEMS
4. OPTIONS & CONSEQUENCES
5. ASSESSMENT OF THE DECISION & ACTION

HUMAN COGNITIVE BIASES



https://www.visualcapitalist.com/wp-content/uploads/2021/08/all-188-cognitive-biases.html



Douglas R. Hofstadter: Gödel, Escher, Bach: An Eternal Golden Braid, p.79 Fig. 18

LIMITS OF NORMATIVE/ AXIOMATIC SYSTEMS

Normative (rule-based) systems have a finite range
Distinguishing true from false is not always simple.

BEYOND COMPLIANCE AND SIMPLE SET OF RULES

(1) THE MEANINGS AND SCOPES OF ETHICS

Ethics of a specific domains (AI Ethics, Healthcare Ethics, Business Ethics, Design Ethics, Engineering Ethics, Research Ethics, etc.

Ethics at a specific level (Theoreticsl Ethics, Practical Ethics, Applied Ethics, Meta-Ethics,…)

Aspects of ethics: Cognitive, Logical, Organizational, Socio-political…

ETHICS AS A PHENOMENON

"**ETHICS THEATHER**" = simulacrum of "ethicity" as a ritual without genuine interest or engagement. Often used for entirely different purposes. Similar to Covid-19 "hygiene theather" like disinfecting streets and houses from the air

Related to:

"Ethical whitewashing" (or "ethical window dressing") - the practice of making 'misleading claims … or implementing superficial measures' in order to defend one's practice and thereby make it appear more ethical than it is in reality (Floridi 2019: 186).

Floridi, L. (2019). Translating principles into practices of digital ethics: Five risks of being unethical. Philosophy & Technology 32:2, 185-193. https://doi.org/10.1007/s13347-019-00354-x

(2) PROACTIVE MULTIPERSPECTIVE RESEARCH ETHICS
ANTICIPATION, LEARNING, RESPECT & TRUST

- Ethics of different research fields – what research fields are addressing and how they are doing that – Information Ethics, Computing ethics, Bioethics, Medical ethics
- Ethics of researchers – codes of conduct – virtues and utility http://ec.europa.eu/research/participants/data/ref/fp7/89888/ethics-for-researchers_en.pdf
- Study of research ethics itself – can cognitive science help us understand ethics better? http://www.iep.utm.edu/m-cog-sc/
- Can Ethics be defended by unethical means?

(3) BEYOND HUMAN DECISION-MAKING:
DELEGATING RESPONSIBILITIES TO AUTONOMOUS AIs

- AI assistants as autonomous intelligent agents
- With autonomy and intelligence comes the necessity of ethics by design.*
- Artificial Intelligence needs ART (Accountability, Responsibility, Transparency)
- Two-step process proposal sub-symbolic level based on deep learning which in the next step is analyzed by a symbolic-level AI that can compare it with the regulations, rules and best practices.

Dodig-Crnkovic G. and Çürüklü B., Robots - Ethical by Design, Ethics and Information Technology 2011, Volume 14, Number 1, pp. 61-71. https://dl.acm.org/doi/abs/10.1007/s10676-011-9278-2, https://link.springer.com/article/10.1007/s10676-011-9278-2,

4) AI-GENERATED WEB CONTENT:
AUTONOMOUS AI DECISION-MAKING IN THE BACKGROUND AND INFRASTRUCTURE

- Articles and Blog Posts
- Product Descriptions
- News Articles, Automated Journalism, Data Journalism
- Social Media Posts
- Video and Audio Scripts and Captions
- Chatbots and Customer Support

US elections: “Perplexity dove into real-time election tracking” – with additional guardrails to avoid fabrication (“hallucination”) and partial information/misinformation. ChatGPT refused to comment elections.

In the first step, AI is trained on the content produced by humans.

In the next step, an increasing amount of content will be produced by AI and used to train AI.

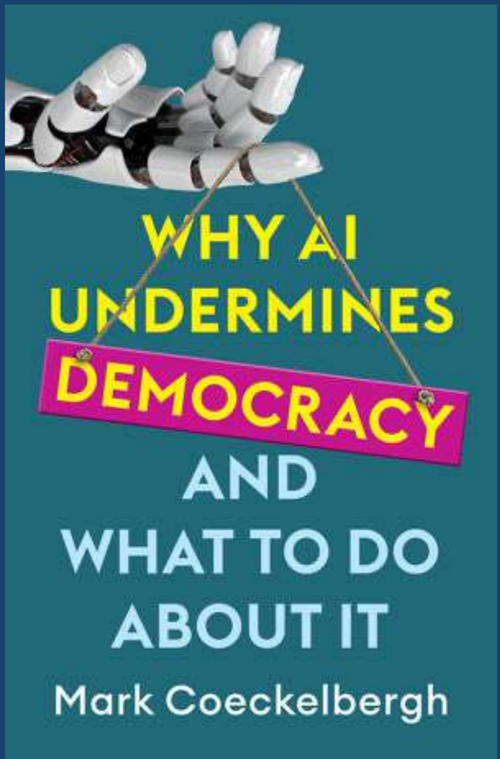
AI ethics – beyond piecemeal evaluations: comprehensive threat surface analysis for advanced AI systems – Upcoming Chalmers AI Ethics Seminar 26/11

"Current AI risk assessment methods often fall short in addressing the full spectrum of potential threats posed by advanced AI systems. This talk will explore the need for a paradigm shift towards comprehensive threat surface analysis, drawing on techniques like Probabilistic Risk Assessment (PRA)."

Anna Katariina Wisakanto, the Center for AI Risk Management & Alignment (CARMA)

Reference: https://link.springer.com/content/pdf/10.1038/s41598-023-34622-w.pdf?pdf=core Bernd Carsten Stahl- Embedding responsibility in intelligent systems: from AI ethics to responsible AI ecosystems

LITERATURE



Mark Coeckelbergh (2022) The Political Philosophy of AI: An Introduction
ISBN: 978-1-509-54855-2

The Political Philosophy of AI: An Introduction

Table of Contents

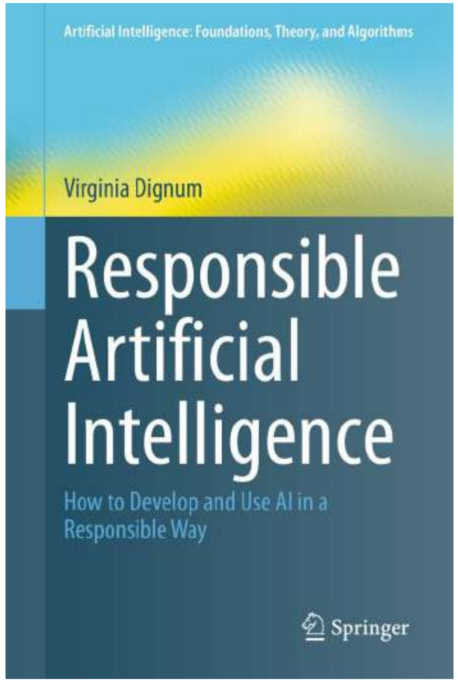
RESPONSIBLE AI: WHY CARE?

- AI systems are designed to act autonomously in our world
- Eventually, AI systems will make *better* decisions than humans in specific well-defined domains

AI is designed by humans (currently), it is an artefact

- "We need to be sure that the purpose put into the machine is the purpose which we really want" (Alignment problem)

Norbert Wiener, 1960 (Stuart Russell) King Midas, c540 BCE



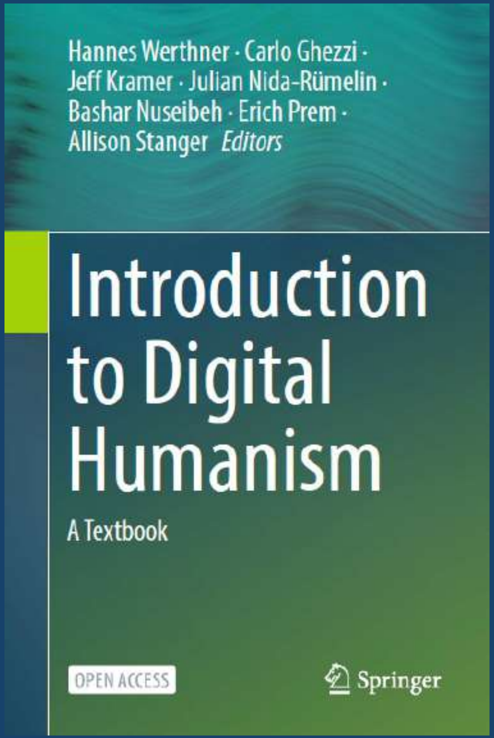
Based on: Responsible Artificial Intelligence, Virginia Dignum, https://www.youtube.com/watch?v=BqwVRzKVz30

Perspectives on Digital Humanism – Open Access

Hannes Werthner, Erich Prem, Edward A. Lee, and Carlo Ghezzi (eds): **Perspectives on Digital Humanism**, Springer, 2022. https://link.springer.com/book/10.1007/978-3-030-86144-5

Introduction to Digital Humanism – A Textbook

Open Access



Hannes Werthner, Carlo Ghezzi, Jeff Kramer, Julian Nida-Rümelin, Bashar Nuseibeh, Erich Prem, and Allison Stanger (eds): **Introduction to Digital Humanism**, Springer, 2024. https://link.springer.com/book/10.1007/978-3-030-86144-5

The Digital Humanism

Digital Humanism is a mindset, a philosophy, a political driving force, a scientific approach, and most of all – a promise to society. That we are determined to build, regulate and develop technology for people, for a better future.

https://informatics.tuwien.ac.at/digital-humanism/
https://informatics.tuwien.ac.at/digital-humanism/#latest-news
https://dighum.ec.tuwien.ac.at/ Digital Humanism movement web page @ TUW – Technical University in Vienna
https://www.youtube.com/@DigitalHumanism/featured?app=desktop Digital Humanism YouTube channel

http://www.gordana.se/work/PRESENTATIONS-files/presentations-pdf/20250108-DIGITAL%20HUMANISM.pdf Digital Humanism lecture


https://www.youtube.com/watch?v=Mccpq8fpUI8
Video Lecture at Carlos III University, Madrid
10 October 2025, Croatian Competence Center for High Performance
Navigating the White-Water
with Digital Humanism
Emergent Intelligent Technologies between Utopia an
CHALMERS
26th International Conference on
Model Driven Engineering Languages and Systems
MODELS 23
1-6 October 20
Västerås, Swed
Navigating the White-Water W
with Digital Humanism
Emergent Digital Technologies between Utopia and Dystopia
na Dodig Crnkovic
Università
della
Svizzera
italiana
21 March 2024/
ity of Italian Switzerland, or University of Lugano
Navigating the White-Water W
ith Digital Humanism
ergent Intelligent Technologies between Utopia and D
na Dodig Crnkovic
Navigating the White-Water W
with Digital Humanism
Emergent Intelligent Technologies between Utopia and Dys
Gordana Dodig Crnkovic
enior Professor of Computer Science at Mälardalen University and
ofessor of Interaction Design, Chalmers University of Technology, Sweden, http://gordana.se/

AI IS LIKE A FIRE-JUST MUCH MORE POWERFUL

We are not only talking of AI as it looks like at this moment. We think of its developments in coming years.

“It is unworthy of excellent (wo)men to lose hours like slaves in the labour of calculation which could safely be relegated to anyone else if machines were used.”

Gottfried Wilhelm Leibniz, Describing, in 1685, the value to astronomers of the hand-cranked calculating machine he had invented in 1673.

REFERENCES

https://deepmind.com/about/ethics-and-society GOOGLE DEEP MIND Ethics & SOCIETY

https://framtidsprao.trr.se/documents/Framtidens_arbetsliv_rapport_WEB.pdf

https://www.youtube.com/watch?v=RXCqKwMHpb0 Ethics of AI @ NYU: Opening & General Issues (1:23:30 - Yann LeCun "Should We Fear Future AI Systems?")

https://www.youtube.com/watch?v=1oeoosMrJz4 AI ethics and AI risk - Ten challenges

https://futureoflife.org/ai-principles/ Asilomar Principles

https://www.microsoft.com/en-us/ai/ai-for-good AI for Earth, AccessibilityHumanitarian Action, Cultural Heritage

https://www.partnershiponai.org PARTNERSHIP ON AI to benefit humanity
Started by Microsoft, Amazon, Google, Facebook, IBM, and Google-owned DeepMind. 2019: 90+ partners, >50% non-profit, 13 countries

https://link.springer.com/book/10.1007/978-3-030-69978-9 Bernd Carsten Stahl- Artificial Intelligence for a Better Future – Open access

https://link.springer.com/content/pdf/10.1038/s41598-023-34622-w.pdf?pdf=core Bernd Carsten Stahl- Embedding responsibility in intelligent systems: from AI ethics to responsible AI ecosystems

REFERENCES

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.

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.

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

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

REFERENCES

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

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

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

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 https://www.quantamagazine.org/artificial-intelligence-will-do-what-we-ask-thats-a-problem-20200130/ Artificial Intelligence Will Do What We Ask. That's a Problem

http://www.gordana.se/work/presentations.html