

Philosophy of Technology and Ethics of AI: From Alan Turing to the Age of Algorithms

Lecturer: Dr. John Dorsch (he/him | they/them)

Course Description:

Artificial Intelligence (AI) is reshaping nearly every aspect of contemporary life—from how we work and communicate to how we understand creativity, responsibility, and even intelligence itself. This course introduces students from all disciplines to the ethical, social, and conceptual challenges posed by AI and modern technology. It combines perspectives from philosophy, media studies, and contemporary journalism to help students think critically about how these technologies influence both society and human self-understanding.

In the first third of the course, we will trace the roots of AI through Alan Turing's classic question, "Can machines think?" and explore how early conceptions of intelligence continue to shape today's debates. We will then turn to Shannon Vallor's *The AI Mirror* to consider how metaphors of mind and intelligence shape our ethical imagination, and to Karen Hao's *Empire of AI* to examine the real-world conditions of AI production—its social inequalities, labor practices, and environmental costs. The remainder of the course focuses on contemporary ethical issues: trust and responsibility in AI systems, the role of bias and transparency, and how technology can both empower and constrain human agency. Weekly readings will include a mix of academic and public-facing texts.

Each weekly theme will be explored in two 90-minute sessions. The first session, led by the instructor, will introduce the topic through a close reading and discussion of the assigned text. Students are not required to have completed the reading in advance of this session. The second session will be student-led, featuring a presentation and critical analysis of the text, followed by an open discussion of its central ideas and implications.

Course Prerequisites:

No prior background in philosophy or computer science is required—only curiosity and a willingness to question how technology is transforming the world we live in.

Course Objectives/Learning Outcomes:

- Develop a deep understanding of the ethical considerations surrounding AI and its impact on society.
- Acquire the ability to critically analyze philosophical texts related to AI, technology, and their interplay.
- Learn the fundamental concepts in the philosophy of technology, including the nature and influence of technology on our world.
- Cultivate the skill of effectively synthesizing and presenting complex philosophical ideas to further the academic discourse.
- Assess the moral and societal implications of AI technology and develop an understanding of ethical theories around its accountability.
- Examine real-world AI implementations to evaluate the practical implications of AI ethics.
- Identify contemporary issues in technology ethics and AI ethics and acquire an understanding of evolving ethical debates.
- Hone critical thinking skills for well-reasoned, practically informed ethical arguments in the context of AI and technology.

Grading Policies:

Mandatory Completion Policy: Note that all mandatory assignments and exams must be completed to the best of your ability in order for your final grade to be issued. Failure to complete a mandatory assignment or exam may result in a failing grade.

Late-Work Policy: Late submissions are generally not accepted. Exceptions will be made only with a valid doctor's note.

Make-Up Work: Make-up work may be allowed in exceptional circumstances, but it requires both instructor permission and a valid doctor's note or similar documents. Students must communicate with the instructor as soon as possible in case of any necessary make-up work.

Attendance Policy: Attendance is mandatory for all AEP, Erasmus, and Czech students, with limited allowances for unexcused absences (typically one per semester). Students who are unwell must contact the AEP academic coordinator to request an excused absence. A doctor's note is required for medical-related absences. Unexcused absences will impact the student's final grade as outlined in individual course syllabi.

Letter Grade	Percentage	Description
A+	97-100	Excellent Work
A	93-97	Outstanding Work
A-	90-92	
B+	87-89	Good work
B	83-86	
B-	80-82	
C+	77-79	Acceptable Work
C	73-76	
C-	70-72	
D+	67-69	Work that is significantly below average
D	63-66	
D-	60-62	
F	0-59	Work that does not meet the minimum standards for passing the course

AEP Academic Integrity Policy

Plagiarism and other forms of academic dishonesty are not tolerated. The use of Artificial Intelligence (AI) for the development of knowledge and learning is encouraged at many stages of the learning process. While we value technology for educational purposes, we also value originality and the retainment of knowledge, and thus using AI for assignments and examinations, even if rephrased, is strictly prohibited and considered an academic integrity violation, unless the instructor explicitly allows for it in the context of evaluated work

AEP Non-Discrimination/Harassment Policy

The AEP program in Prague promotes a diverse learning environment where the dignity, worth, and differences of each individual are valued and respected. Discrimination and harassment, whether based on a person's race, gender, sexual orientation, color, religion, national origin, age, disability, or other legally protected characteristics, are repugnant and completely inconsistent with our objectives. Retaliation against individuals for raising good faith claims of harassment and/or discrimination is prohibited.

AEP Diversity Policy

AEP is committed to fostering an inclusive and welcoming community that values diversity in all its forms. We believe that one of the most meaningful lessons of studying abroad is learning to navigate and appreciate differences with curiosity and an open mind. While engaging across differences can sometimes be challenging or uncomfortable, these moments are essential for growth and learning. We recognize that every member of our community, even with the best intentions, may occasionally make missteps. Our commitment is to provide a supportive environment where respectful and honest dialogue helps us learn from these experiences, ensuring that every student has the opportunity to thrive and broaden their perspective.

Learning Resources

- **Required Readings:** Weekly journal articles and book chapters will be assigned as mandatory readings, covering various topics related to technology and AI ethics. These readings will form the core material for the course.
- **Course Homepage:** The Moodle homepage will serve as the central hub for accessing all the required materials. Students will find links to the assigned readings, lecture notes, and supplementary resources on this platform here.
- **Lecture Notes:** Select lecture notes and presentation slides provided by the instructor will offer additional insights and explanations of the course content.
- **Philosophy Encyclopedia:** If you would like to read up on a topic in philosophy, the best place to start is the [Stanford Encyclopedia of Philosophy](#), which is written by experts and intended to serve as a general introduction and counts as an excellent reference for submitted materials, like essays.
- **Machine Learning Glossary:** If you need to look up AI or machine learning terms, I recommend [Google's Machine Learning Glossary](#).

Assessment

The assessment for 3 US credits consists of attendance and active class participation (20%), a 3-page essay (40%), and a presentation (40%).

Essay:

Grading: Essays are 40% of the grade, worth 100 possible points, consisting of 5 parts. Please see below. This grading system intentionally reserves a broad upper band (worth 30 points) for strong and excellent work. By allowing some margin within the top range, the grading approach recognizes that progress in philosophical writing involves not only improving clarity, precision, and argumentative structure but also learning how to advance well-supported claims. The goal is to give students room to grow into the skill of writing a successful philosophical essay — a form of writing that takes practice, reflection, and sustained effort to master.

Essay's Total Points	US Grade
70-100	90-100
60-69	80-89
50-59	70-79
40-49	60-69
Below 40	Below 60

- Introduction: 50-100 words 10 points
- Exposition: 100-300 words 40 points
- *Argument: 750-1000 words 40 points
- Conclusion: 50-100 words 10 points
- Reference Page: Points can be lost for mistakes.

* Students often lose points in the Argument section because they present claims without fully supporting them, or because the structure of their reasoning is unclear. A strong philosophical argument requires not only stating a position but also **defending it with reasons**, considering possible objections, and showing why those objections do not undermine the main claim. Essays that remain too descriptive or rely on assertion rather than justification tend to fall short in this category.

Length: 1250 - 1500 words or 3 pages, and it must not exceed this limit.

Submission Date: The essay must be submitted via email to dorsch@cz.flu.cz by a date to be announced at the start of the course, and it must not be submitted later.

Topic: Students choose one question from 3 possible questions to write their essay about. Possible questions will be announced in the seminar. Students may write about a self-selected topic, but they must discuss the topic with the course instructor.

Content: The course instructor has provided an extensive outline of what is expected in your essay. Please see the course's Moodle page.

Formatting: Non-anonymized, in English (or German), and in APA format (i.e., Times New Roman 12, 1.5 line spacing, justified alignment, etc.).

Citations: (a) In-text citations should be used, and (b) a reference page is required.

(a) In-text citations should appear as follows: This response thus posits two classes of spatial representations: allocentric and egocentric (Klatzky, 1998).

(b) The reference page should be formatted as follows (with hanging indent 0.5 cm):

Klatzky, R. (1998). Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections. *Spatial Cognition*, 1404, pp.1-17.[https://doi.org/ 10.1007/3-540-69342-4_1](https://doi.org/10.1007/3-540-69342-4_1).

Presentation

Grading: 40% of the final grade.

Content: Presentations will be graded along the following three factors: I. Analysis, II. Critique, and III. Active Teaching. Presentations can be done individually or in a group, but if in a group, the contribution of each presenter must be clearly identifiable. You must sign up on the shared Google Doc by the third week on the course's Moodle page. Please send presentation slides to dorsch@cz.flu.cas for them to be uploaded.

- **I - Analysis in the Presentation (10-20 minutes)**

- What is the central point or claim being made? Summarize it accurately in your own words.
- What kind of point is it? For instance, is it explanatory, descriptive, interpretive, or argumentative?
- What kind of support is offered? Identify whether the author relies on data, examples, testimony, conceptual analysis, or narrative evidence.
- How is the reasoning or presentation structured? Outline how the author moves from one idea to the next and how the main point is developed or defended.
- What assumptions or perspectives shape the discussion? Note any underlying viewpoints, values, or contextual factors that influence how the material is presented.

- **II - Critique in the Presentation (10-20 minutes)**

- How convincing is the reasoning or presentation? Does the author's conclusion follow logically or coherently from the material presented?
- Are the sources or forms of support reliable and relevant? Consider whether the evidence, examples, or reasoning genuinely support the main point.
- Are there alternative interpretations or explanations? Identify other ways the issue, data, or argument might be understood or approached.
- What limitations or gaps can be identified? Note where the reasoning, evidence, or framing might be incomplete, biased, or open to question.

- **III - Active Teaching during and after the Presentation (45-60 minutes)**
 - **Content:** Presentations should be accompanied by at least two groupwork tasks. Choose a combination of the following:
 - **Discussion tasks:** These always consist of two parts. First, short discussions will take place in small groups. Second, a large group discussion in the seminar will refer to the entirety of the group discussions.
 - **Discussion with Idea Collection:** Presenters must first select the topic for discussion (20-45 minutes).
 - **Text Analysis and Interpretation:** Different sections of the reading will be selected by the presenters and distributed to the groups (20-45 minutes).
 - **Justification Analysis:** Presenters must first present the justification (or justifications) before the seminar is divided into groups (10-20 minutes).
 - **Study Analysis:** Groups will either interpret the same study OR different studies, and presenters decide who gets what (20-45 minutes).
 - **Roundtable tasks:** These consist of only one part. Questions are asked in a round (one by one), and everyone must contribute.
 - **Initial Thoughts?** In the round, participants will be asked about their initial thoughts on the weekly reading, and each person must contribute (5-10 minutes).
 - **Summarize and Add:** “;He/she said... and I add...”; This is not a memory game! You only need to summarize the ideas of a previous person, and you can choose that person yourself (10-20 minutes).
 - **Other Tasks**
 - **Survey:** Create a survey for free using Google Forms:
<https://www.google.com/forms/>; Create a QR code for it for free here:
<https://www.qrcode-generator.de> (5-10 minutes)
 - **Debate:** The seminar is divided into two groups, and a discussion question is posed. One group is the “yes/for” group, and the other is the “;no/against” group. Give both groups 10 minutes of preparation time. The presenters are responsible for moderating the debate and determining which group won.

Course Policies

Statement about Student Wellness: As a student, you may experience a range of challenges that can interfere with learning, such as strained relationships, increased anxiety, substance use, feeling down, difficulty concentrating and/or lack of motivation. These mental health concerns or stressful events may diminish your academic performance and/or reduce your ability to participate in daily activities. As your teacher, I am here to support you.

Mobile devices: Mobile devices are welcome in the course but should be used for educational purposes only in order to minimize distractions.

Weekly Schedule

Week 1

CEE Introductory Lecture Series

AEP Introductory Lecture Series

Week 2

Initial Considerations: What is AI Ethics?

Gain a basic understanding of the key issues in philosophy of technology and the ethical questions surrounding AI.

Reading

- None (first day of seminar)

Week 3

Theory and History

Explore the historical development of philosophical perspectives on technology and its impact on society.

Readings

- Turing (1950) "Can Machines Think?"

Week 4

Contemporary Introduction

Delve into the metaphors used to describe AI and how these shape our thinking

Readings

- Vallor (2024) "The AI Mirror" (Chapter 1)

Week 5

Critical Analysis

Continue the analysis of metaphors to describe AI, and focus on potential societal implications.

Readings

- Vallor (2024) “The Empathy Box” from her book “The AI Mirror” (Chapter 5).

Week 6

Practical and Societal Implications

Assess the societal and ecological impact of AI.

Readings

- Hao’s (2025) “Plundered Earth” from her book “Empire of AI” (Chapter 12)

Week 7

Self-Study and Self-Assessment with Mid-Term Exam

Week 8

Political Philosophy

Investigate the concept of “algocracy” and its implications for political freedom.

Readings

- Danaher (2020) “Freedom in the Age of Algocracy”

Week 9

Freedom and Identity

Analyze the ethical dimensions of AI with respect to race and gender.

Readings

- Gebru (2020) “Race and Gender”

Week 10

Metaphysics and Mind

Explore the considerations surrounding the concept of AI “artificial persons”.

Readings

- Schwitzgebel & Garza (2020) “Designing AI with Rights, Consciousness, Self- Respect, and Freedom”

Week 11

Debate over Trustworthy AI – Part 1.

Explore the notion of trustworthy AI and its attribution of responsibility.

Readings

- Dorsch & Deroy (2024) “Why We Don’t Need Morally Trustworthy AI”

Week 12

Debate over Trustworthy AI –Part 2.

Analyze critiques and responses in reference to the “Trustworthy AI” Debate.

Readings

- Franke’s (2024) “The Limits of Calibration”
- Dorsch, Deroy, and Moll (forthcoming) “Why We Need Epistemically, not Morally, Trustworthy AI”

Week 13

Course Overview and Final Essay Preparation