



SUBJECT DATASHEET

Artificial intelligence and ethics

BMEGT41V105

I. SUBJECT DESCRIPTION

1. SUBJECT DATA

Subject name

Artificial intelligence and ethics

ID (subject code)

BMEGT41V105

Type of subject

contact lessons

Course types and lessons

<i>Type</i>	<i>Lessons</i>
Lecture	2
Practice	0
Laboratory	0

Type of

assessment

seminar grade

Number of

credits

2

Subject Coordinator

<i>Name</i>	<i>Position</i>	<i>Contact details</i>
-------------	-----------------	------------------------

Dr. Héder Mihály	associate professor	heder.mihaly@gtk.bme.hu
------------------	---------------------	-------------------------

Educational organisational unit for the subject

Department of Philosophy and History of Science

Subject website

<https://edu.gtk.bme.hu>

Language of the subject

magyar - HU

Curricular role of the subject, recommended number of terms

Programme: **Any programme**

Subject Role: **Elective**

Recommended semester: **0**

Direct prerequisites

Strong None

Weak None

Parallel None

Exclusion None

Validity of the Subject Description

Approved by the Faculty Board of Faculty of Economic and Social Sciences, Decree No: 580884/8/2023 registration number. Valid from: 29.11.2023.

2. OBJECTIVES AND LEARNING OUTCOMES

Objectives

The main objective of the subject is to map the ethical problems related to artificial intelligence and automatization and to show the possible solutions to these problems. The students get a glimpse into the contemporary situation of AI in terms of technological preparedness and market penetration. The most important ethical frameworks are discussed. While the course addresses both topics in detail, it does not have pre-requisites as it was designed to be accessible for all students of a technical university. At the second part of the semester, the two main topics get synthesized with a particular view to the recent findings and proposals of the related international engineering organizations (IEEE, ACM, VDI, university centres). The problems are framed by using case studies (e.g., ethical concerns of autonomous cars, recommendation systems, ethical problems of browsers with AI, unemployment and AI, AI and issues of world security, rights of AI agents). The examples might change to represent the actual trending topics of the field

Academic results

Knowledge

1. the conventional terminology of AI.
2. the conventional terminology of ethics.
3. the ethical concerns related to AI.
4. various approaches of ethics related to the cases covered in the subject.
5. how the discipline being studied fits into a larger disciplinary and social scheme, how it relates to the neighbouring fields, how these fittings give rise to using certain contexts.
6. the fundamentals of various mechanisms of social decision-making.

Skills

1. to use professional terms, basic concepts and advanced elements of the special terminology of the profession.
2. to see and treat science in its embeddedness, to use an interdisciplinary approach.
3. to analyze, evaluate, draw conclusions and synthesize explanations in their professional work.
4. to provide a critical analysis of information by using a wide array of well-founded techniques.
5. to participate in lifelong learning.
6. to use a multi-sided, interdisciplinary approach to identify specific problems, to explore and define the detailed theoretical and practical background that is needed to solve these problems.
7. to discover facts and basic relationships by applying the theories and methods learned. To systematize, to analyze, to draw conclusions, to give critical remarks,
8. to prepare proposals of decision-making and to make decisions both in a routine and in a non-routine - domestic and international - environment.

Attitude

1. Accepts and consistently and plausibly represents the diversity of the perspectives of social sciences in the related narrower and wider environment.
2. Demonstrates an open mind to critical self-evaluation, to various forms of training, to the self-help forms of intellectual worldview.
3. Endeavours for self-development in these areas.
4. Has problem-centric perspective and problem-solving thinking.

Independence and responsibility

1. - On the basis of professional guidance, thinks through broad and specific professional issues and
2. based on specific sources.
3. - Independently performs his/her work with critical evaluation and continuous correction of his/her work.
4. - Participates responsibly in the development and justification of professional views.
5. - Assumes responsibility for the views that underpin his/her area of expertise.

Teaching methodology

Lectures, optional individual or group exercises.

Materials supporting learning

- Stuart, Russel, and P. Norvig. „Mesterséges Intelligencia modern megközelítésben.
- Boros Gábor (szerk). „Filozófia.
- Diák és jegyzetek a www.filozofia.bme.hu oldalon. Slides and notes on the official site of the subject.
- Gerdes, J. C., & Thornton, S. M. Implementable ethics for autonomous vehicles. In *Autonomes fahren* (pp. 87- 102). Springer Vieweg, Berlin, Heidelberg, (2015).
- MIT OpenCourseWare: Engineering Ethics
- <https://ocw.mit.edu/courses/engineering-systems-division/esd-932-engineering-ethics-spring-2006/>
- Molnár László, Environmental Ethics and Engineering Ethics Courses at the BUTE: What Kind of Ethics Do We Teach? In: Banse, Gerhard, Imre Hronszky, and Gordon L. Nelson, eds. *Rationality in an uncertain world*. Vol. 5. edition sigma, (2005).
- Letölthető anyagok/Downloadable contents:
- IEEE Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems
- <https://ethicsinaction.ieee.org/>
- European Commission: Ethics guidelines for trustworthy AI:
- https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=58477

- Magyar Mérnöki Kamara Országos Etikai-fegyelmi Szabályzat
- <https://mmk.hu/tudastar/dokumentumtar/szabalyzatok/etikai-fegyelmi-szabalyzat.pdf>
- Verein Deutscher Ingenieure Ethische Grundsätze des Ingenieurberufs
- <https://www.vdi.de/bildung/ethische-grundsaeetze/ethische-grundsaeetze/>
- IEEE Code of Ethics
- <https://www.ieee.org/about/corporate/governance/p7-8.html>

II. SUBJECT REQUIREMENTS

TESTING AND ASSESSMENT OF LEARNING PERFORMANCE

General Rules

A 2.2. pontban megfogalmazott tanulási eredmények értékelése: kettő évközi írásbeli teljesítménymérés és az aktív órai részvétel alapján (opcionális) történik.

Performance assessment methods

A. Összegző tanulmányi teljesítményértékelés: a tantárgy és tudás, képesség típusú kompetenciaelemeinek komplex, írásos értékelési módja zárthelyi dolgozat formájában. A dolgozat tudáselemekre, értelmezési feladatokra, következtetési feladatokra koncentrált írásos kérdésválaszolás formájában. munkaidő anyagrésztől függően 30-60 perc. Minden összegző tanulmányi teljesítményértékelésen legalább a szerezhető pontok 50%-át el kell érni a tárgy teljesítéséhez. B. Részteljesítmény-értékelés (órai aktív részvétel, opcionális): a tantárgy tudás, képesség, attitűd, valamint önállóság és felelősség típusú kompetenciaelemeinek egyszerűsített értékelési módja, melynek megjelenési formája a felkészült megjelenés és tevékeny részvétel a gyakorlat folyamatában, felkérésre vezetett példamegoldás a hallgatók előtt; az egységes értékelési elveket a tantárgyfelelős és a tantárgy előadója együttesen határozza meg.

Percentage of performance assessments, conducted during the study period, within the rating

- 1. összegző tanulmányi teljesítményértékelés : 50
- 2. összegző tanulmányi teljesítményértékelés : 50
- opcionális: 50
- összesen: 100

Percentage of exam elements within the rating

Conditions for obtaining a signature, validity of the signature

Issuing grades

Excellent	90
Very good	81–90
Good	71–80
Satisfactory	61–70
Pass	50–60
Fail	49

Retake and late completion

Javítás esetén a korábbi és az új eredmény közül a pótláskor vagy javításkor elért eredmény kerül beszámításba

Coursework required for the completion of the subject

részvétel a kontakt tanórákon	28
felkészülés a teljesítményértékelésekre	16
házi feladat elkészítése	0
kijelölt írásos tananyag önálló elsajátítása	16
vizsgafelkészülés	0
összesen	60

Approval and validity of subject requirements

Consulted with the Faculty Student Representative Committee, approved by the Vice Dean for Education, valid from: 06.11.2023.

III. COURSE CURRICULUM

THEMATIC UNITS AND FURTHER DETAILS

Topics covered during the term

1) Az MI Etika problématerképe 2) Az MI és a Munka jövője 3) Az MI és a Technológiába zártság 4) Gépi Etika: transzparencia 5) Gépi Etika: megmagyarázhatóság (XAI) 6) Gépi Etika: méltányosság és részrehajlás algoritmusokban (bias; de-bias) 7) Gépi Etika: értékreprezentációk algoritmusokban (value alignment) 8) A mesterséges személy morális és társadalmi státusza: attribútum alapú megközelítések 9) A mesterséges személy morális és társadalmi státusza: reláció alapú megközelítések 10) Az EU AI törvény filozófiai háttere 11) Az IEEE P7001 szabványtervezet 12) Az IEEE 7000-2021 szabványtervezet

Additional lecturers

Dr. Danka István egyetemi adjunktus danka.istvan@gtk.bme.hu

Dr. Ziegler Zsolt egyetemi adjunktus ziegler.zsolt@gtk.bme.hu

Approval and validity of subject requirements