



TANTÁRGYI ADATLAP

A MESTERSÉGES INTELLIGENCIA ETIKÁJA

BMEGT41V105

I. TANTÁRGYLEÍRÁS

1. ALAPADATOK

Tantárgy neve

A MESTERSÉGES INTELLIGENCIA ETIKÁJA

Azonosító

BMEGT41V105

A tantárgy jellege

kontaktórás tanegység

Kurzustípusok és óraszámok

<i>Típus</i>	<i>óraszám</i>
Előadás	2
Gyakorlat	0
Laboratórium	0

Tanulmányi

teljesítményértékelés
(minőségértékelés)

típusa

félévközi
érdemjegy

Kreditszám

2

Tantárgyfelelős

<i>Neve</i>	<i>Beosztása</i>	<i>Email címe</i>
-------------	------------------	-------------------

Dr. Héder Mihály	egyetemi docens	heder.mihaly@gtk.bme.hu
------------------	-----------------	-------------------------

Tantárgyat gondozó oktatási szervezeti egység

Filozófia és Tudománytörténet Tanszék

A tantárgy weblapja

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

A tantárgy oktatásának nyelve

magyar - HU

A tantárgy tantervi szerepe, ajánlott féléve

Szak: **Bármely szak**

Tantárgy szerepe: **Szabadon választható**

Ajánlott félév: **0**

Közvetlen előkövetelmények

Erős Nincs

Gyenge Nincs

Párhuzamos Nincs

Kizáró feltételek Nincs

A tantárgyleírás érvényessége

Jóváhagyta a Gazdaság- és Társadalomtudományi Kar Kari Tanácsa (2023.11.29.) az 580884/8/2023 iktatószámon hozott határozatával, amely érvényes 2023.11.29-től.

2. CÉLKITŰZÉSEK ÉS TANULÁSI EREDMÉNYEK

Célkitűzések

A tantárgy célja, hogy feltérképezze a mesterséges intelligencia és az automatizálás által felvetett etikai problémákat és a lehetséges megoldási stratégiákat. A kurzus elején a hallgatók áttekintést kapnak a mesterséges intelligencia aktuális helyzetéről a technológiai készültség és a piaci elterjedtség dimenzióiban, majd a legfontosabb etikai keret- rendszerek is bemutatásra kerülnek. A kurzus mindkét témát kellő mélységben tárgyalja, ám előismeretet nem igényel, így az anyag megértése bármely műegyetemi hallgatótól elvárható. A szemeszter második felében a két téma szintézise történik, különös figyelemmel a nemzetközi mérnök- és interdiszciplináris (pl. IEEE, ACM, VDI, fontosabb egyetemi központok) szervezetek eredményeire és ajánlásaira a témában. A problémateret mindvégig esettanulmányok illusztrálják (pl. önvezető autó etikai kérdései; ajánlórendszerek, intelligens keresők által generált etikai problémák; munkahelyek MI általi megszűnése; MI és világbiztonság, MI ágensek jogai). A példák az aktualitásoknak megfelelően változhatnak.

Tanulmányi eredmények

Tudás

1. Ismeri a mesterséges intelligencia általánosan használt fogalomrendszerét,
2. ismeri az etika általánosan használt fogalomrendszerét,
3. érti a mesterséges intelligencia által felvetett etikai problémákat,
4. ismeri a különféle etikai megközelítéseket a tárgy során tárgyalt problémák megoldására,
5. ismeretei kiterjednek a szakterület tágabb rendszerben való elhelyezésére, a rokon szakterületekhez való kapcsolatok meg- és felismerésére, a tágabb rendszer adta lehetőségek és a hatásrendszerre vonatkozó kontextusok használatára,
6. megfelelő és elegendő ismeretanyaggal rendelkezik ahhoz, hogy eligazodjon a társadalmi döntéshozatal különböző mechanizmusában.

Képesség

1. Szakmai szóhasználata során magabiztosan használja a szakma szókincsét, a szakma szaktudományos alapfogalmait, és a rájuk épülő speciális szakszókészlet elemeit,
2. rendelkezik az új látásmód képességével, képes a tudományt és annak környezetét interdiszciplináris szemlélettel megközelíteni,
3. szakmai feladatainak megoldásában képes önálló elemzésre, értékelésre, és következtetések és magyarázatok szintetizálására,
4. képes az információk kritikus elemzéséhez és feldolgozásához kellően megalapozott technikák széles skálájának alkalmazására,
5. képes az élethosszig tartó tanulás folyamatában való részvételre,
6. sokoldalú, interdiszciplináris megközelítéssel azonosít speciális szakmai problémákat, feltárja és megfogalmazza az azok megoldásához szükséges részletes elméleti és gyakorlati hátteret,
7. a tanult elméletek és módszerek alkalmazásával tényeket és alapvető összefüggéseket tár fel, rendszerez és elemez, önálló következtetéseket, kritikai észrevételeket fogalmaz meg, döntés-előkészítő javaslatokat készít, döntéseket hoz rutin- és részben ismeretlen - hazai, illetve nemzetközi - környezetben is,
8. képes gondolatait rendezett formában szóban és írásban kifejezni.

Attitűd

1. Együttműködik az ismeretek bővítése során az oktatóval és hallgató társaival,
2. folyamatos ismeretszerzéssel bővíti tudását,
3. nyitott a kritikus önértékelésre, a szakmai továbbképzés különböző formáira, az értelmiségi világlátás önfejlesztő módszereire és törekszik önmaga fejlesztésére e területeken,
4. problémacentrikus látásmóddal, problémamegoldó gondolkodással rendelkezik.

Önállóság és felelősség

1. - Szakmai útmutatás alapján végzi átfogó és speciális szakmai kérdések végiggondolását és adott források alapján történő kidolgozását.
2. - Önállóan végzi munkáját tevékenysége kritikus értékelése és folyamatos korrekciója mellett.
3. - Felelősséggel részt vállal szakmai nézetek kialakításában, indoklásában.
4. - A szakterülete megalapozó nézeteit felelősséggel vállalja.

Oktatásmódszertan

Előadások, opcionális önállóan és csoportmunkában készített feladatok.

Tanulástámogató anyagok

- 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. TANTÁRGYKÖVETELMÉNYEK

A TANULMÁNYI TELJESÍTMÉNY ELLENŐRZÉSE ÉS ÉRTÉKELÉSE

Általános szabályok

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

Teljesítményértékelési módszerek

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.

Szorgalmi időszakban végzett teljesítményértékelések részaránya a minősítésben

- 1. összegző tanulmányi teljesítményértékelés : 50
- 2. összegző tanulmányi teljesítményértékelés : 50
- részteljesítmény értékelés (aktív részvétel) : 50
- összesen: 100

Vizsgaelemek részaránya a minősítésben

Az aláírás megszerzésének feltétele, az aláírás érvényessége

Érdemjegy-megállapítás

Jeles	90
Jeles	81–90
Jó	71–80
Közepes	61–70
Elégséges	50–60
Elégtelen	49

Javítás és pótlás

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

A tantárgy elvégzéséhez szükséges tanulmányi munka

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

A tantárgykövetelmények jóváhagyása és érvényessége

A Kari Hallgatói Képviselőt véleményezése után jóváhagyta dr. Lógó Emma oktatási dékánhelyettes 2023.11.06-án. Érvényes 2023.11.06-tól.

III. RÉSZLETES TANTÁRGYTEMATIKA

TEMATIKAI EGYSÉGEK ÉS TOVÁBBI RÉSZLETEK

A félévben sorra vett témák

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

További oktatók

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

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

A tantárgykövetelmények jóváhagyása és érvényessége

A Tantárgyi adatlap I. és II. részén túli III. részét az érintett szak(ok) szakfelelőse(i)vel való egyeztetés alapján az 1.8. pontban megjelölt Filozófia és Tudománytörténet Tanszék vezetője hagyja jóvá.