Towards a Philosophy of Artificial Intelligence

A mesterséges intelligencia filozófiája felé

ABSTRACTS

RÖVIDLETEK

Ed. by Kristóf Nyíri / Szerkesztette Nyíri Kristóf

2026

Towards a Philosophy of Artificial Intelligence

Join Zoom Meeting:
https://us02web.zoom.us/j/88537975328?pwd=b9CgQLdK8lrFrOKWTamIbtUAU7PG71.1
Meeting ID: 885 3797 5328
Passcode: 982546

TOWARDS A PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

Blended physical-online event to be held on Oct. 6, 2026,
organized by the Hungarian Academy of Sciences

The welcoming address will be given by Mihály Pósfai,
President of the Hungarian Academy of Sciences

*** PROGRAM ***

Abstracts

Confirmed invited speakers:
Valéria CSÉPE
András FALUS
John HALDANE
James E. KATZ
Csaba PLÉH
Stuart J. RUSSELL
Barry SMITH
Eörs SZATHMÁRY
Balázs SZEGEDY
Zoltán TANÁCS
Xu WEN

Participants

Having questions? Write to nyirik@gmail.com.

Join Zoom Meeting

https://us02web.zoom.us/j/88537975328?pwd=b9CgQLdK8lrFrOKWTamIbtUAU7PG71.1
Meeting ID: 885 3797 5328
Passcode: 982546

Dress code: informal but not casual.

TOWARDS A PHILOSOPHY OF ARTIFICIAL INTELLIGENCE

Blended physical-online event to be held on Oct. 6, 2026.
13:00–19:00 CET, organized by the Hungarian Academy of Sciences. Site: HAS Main Building, "Nagyterem"

PROGRAM

13:00–13:15: Welcoming address given by Mihály Pósfai,
President of the Hungarian Academy of Sciences
13:15–13:20: Introductory talk by Kristóf Nyíri
13:20–13:45: John Haldane, "Towards a Philosophy of Culture"
13:45–14:10: Valéria Csépe, "Cognitive Offloading:
How to Preserve Human Reasoning
in an AI-dominated World?"
14:10–14:35: Eörs Szathmáry, "How AI is Becoming More Biological"
14:35–15:00: Csaba Pléh, "Relationships between Human Thinking
and Machine Based Solutions:
Remarks of an Outsider"
15:00–15:25: András Falus, "Philosophical-Psychological Perspectives
on Medical AI"
15:25–15:35: Coffee break
15:35–15:40: Stuart Russell, "Provably Beneficial AI: Open Problems"
(video presentation)
15:40–16:05: Xu Wen "Between Dao and Machine:
Towards a Chinese Philosophy of Artificial Intelligence
Rethinking Agency, Harmony, and Human Responsibility
in the Age of AI"
16:05–16:30: James E. Katz, "Saying Thanks to No One:
Buber, Kant, and the Case for Apparatgeist"
16:30–16:55: Barry Smith , "AI Washing"
16:55–17:10: Attila Orbán, "Memory, Tradition and Artificial Intelligence:
Towards an Ontology of Cultural Continuity"
17:10–17:25: Gabriella Manzenreiter, "From Creation to Co-Creation:
Philosophical Perspectives
on Human and Machine Creativity"
17:25–17:35: George Pór, "Can AI Herald a Good Society?"
17:35–17:45: Neville Li, "The New Tower of Babel:
The Modern Attempt to Reach unto Heaven
through the AI Boom"
17:45–18:10: Balázs Szegedy, "On the Mathematics of AI
and Its Philosophical Implications"
18:10–18:30: Zoltán Tanács, "How Politics Should Cope with AI"
18:30–19:00: Concluding discussion

Bilingual booklet. Abstracts / Rövidletek (Conference, Hungarian Academy of Sciences, Oct. 6, 2026 / Konferencia, Magyar Tudományos Akadémia, 2026. okt. 6: Towards a Philosophy of Artificial Intelligence / A mesterséges intelligencia filozófiája felé). Edited by Kristóf Nyíri / Szerkesztette Nyíri Kristóf

Conference speakers / A konferencia előadói: Kristóf Nyíri / Nyíri Kristóf; John Haldane; Valéria Csépe / Csépe Valéria; Eörs Szathmáry / Szathmáry Eörs; Csaba Pléh / Pléh Csaba; András Falus / Falus András; Xu Wen; James E. Katz; Attila Orbán / Orbán Attila; Gabriella Manzenreiter / Manzenreiter Gabriella; George Pór; Neville Chi Hang Li; Balázs Szegedy / Szegedy Balázs; Zoltán Tanács / Tanács Zoltán

This booklet was published with the support of TARR Ltd.

A kiadvány a TARR Kft. támogatásával jelent meg.

ISBN 978-615-02-8098-1

Contents / Tartalom

Kristóf Nyíri, "Introductory talk".

Ladies and Gentlemen, dear Colleagues,

First, let me express my gratitude to Mihály Pósfai, President of the Hungarian Academy of Sciences, for his wonderful welcoming address. Secondly, I wish to thank my co-organizers – above all Attila Orbán, the moderator of today's event – for their invaluable help in bringing this conference into being. And third, and most importantly, let me thank our speakers for their extraordinary contributions. This will be an outstanding conference, one that will leave a lasting impact.

It is not the first conference on AI that I have organized, nor indeed the first one in which I have taken part. The very first was in 1988: the International Summer School *Philosophical Perspectives on Artificial Intelligence* in Bolzano. My lecture, "Wittgenstein and the Problem of Machine Consciousness", was published in 1989. In that talk I argued that it was the famous British mathematician Alan Turing who first pointed out how machine consciousness might come about. I quoted a passage from a 1948 paper of his, and let me quote it again here: "One way of setting about our task of building a 'thinking machine' would be to take a man as a whole and to try to replace all the parts of him by machinery. He would include television cameras, microphones, loudspeakers, wheels and 'handling servo-mechanisms' as well as some sort of 'electronic brain'. … In order the machine should have a chance of finding things out for itself it should be allowed to roam the countryside, and the danger to the ordinary citizen would be serious."

Today it seems that the danger Turing mentions might be even more serious for women than for men. A very recent article in *The Guardian* reports on sex dolls equipped with AI, and describes how some owners mutilate them in the cruellest possible ways, often "killing" them. As the article notes, research on this phenomenon remains inconclusive: "We currently lack robust longitudinal data to definitively prove whether such doll use is protective or, conversely, is a behavioural stepping stone that increases real-world offending

risk. I think we urgently need research on this question as more and more technologies enter the market to simulate human sexuality."

Whatever the danger to women may be, there is another danger that does not require further research, because it is already our reality, dear colleagues – the reality in which we are living right now. AI is part of the climate catastrophe of the spirit. Education, especially higher education, is collapsing. The scientific publication industry is collapsing. Children's psychological health is collapsing.

Is there a way out?

These are the problems we shall discuss at today's event. I wish you a very successful conference.

***

Nyíri Kristóf, „Bevezető hozzászólás".

Hölgyeim és Uraim, kedves kollégák,

Először is hadd fejezzem ki hálámat Pósfai Mihálynak, a Magyar Tudományos Akadémia elnökének, csodálatos köszöntőjéért. Másodszor köszönetet szeretnék mondani társszervezőimnek – mindenekfelett Orbán Attilának, a mai esemény moderátorának – felbecsülhetetlen segítségükért, mely által ez a konferencia létrejöhetett. Harmadszor és legfontosabbként, hadd köszönjem meg előadóinknak rendkívüli kontribúcióikat. Kiemelkedő konferenciára számítok, olyanra, amely maradandó hatást gyakorol.

Nem ez az első MI témájú konferencia melyet szerveztem, sőt nem is az első, amelyen résztvehettem. A legelső, 1988-ban, a *Philosophical Perspectives on Artificial Intelligence* című Nemzetközi Nyári Iskola volt, Bolzano-ban. Előadásom, a „Wittgenstein és a gépi intelligencia problémája", 1989-ben került publikálásra. Abban az előadásban fő állításom úgy hangzott, hogy a híres brit matematikus Alan Turing volt az első, aki rámutatott, hogy hogyan is jöhet létre gépi intelligencia. Idéztem egy bekezdést „Intelligent Machinery" című 1948-as konferencia-írásából, és hadd idézzem azt a bekezdést most újra: „Van egyfajta módja annak, hogy 'gondolkodó gépet' építsünk. Ez volna: tekintsük egy ember egészét, és próbáljuk

minden részét gépre cserélni. Tartalmazna televíziós kamerákat, mikrofonokat, hangszórókat, kerekeket, és 'szervomechanizmusok vezérlését', továbbá valamiféle 'elektronikus agyat'. … Ahhoz, hogy a gépnek esélye legyen arra, hogy a maga számára felfedezzen dolgokat, meg kellene engednünk számára, hogy a vidéken kószáljon, s ezzel a közönséges polgár számára komoly veszély állna elő…"

Ma úgy tűnik, hogy a Turing által említett veszély súlyosabb lehet a nők, mint a férfiak számára. A *The Guardian* egy egészen friss cikke MI-vel ellátott szexbabákról szól, és leírja, hogy azokat némely gazdájuk a lehető legkegyetlenebb módon megcsonkítják, gyakran meg is „ölik". A cikk megjegyzi, hogy az erre a jelenségre vonatkozó kutatás egyelőre nem ad módot következtetések levonására: „Jelenleg nincsen elegendő masszív hosszmetszeti adatunk annak végleges bizonyítására, hogy vajon az ilyen babahasználat védekező jellegű, avagy ellenkezőleg, olyan viselkedésbeli lépcsőfok, amely növeli a bántó magatartás kockázatát a való világban. Úgy gondoljuk, hogy sürgős szükségünk van további kutatásokra, hiszen egyre több az emberi szexualitást szimuláló technológia lép be a piacra."

Bármi legyen is a nőkre leselkedő veszély, van itt másik veszély is, amely viszont nem igényel további kutatást, mivel immár, kedves kollégák, a valóságunk – a valóság, melyben éppen most élünk. Az MI a szellem klímakarasztrófájának része. Az oktatás, különösen a felsőoktatás, összeomlóban. A tudományos publikációs ipar összeomlóban. A gyermekek lelki egészsége összeomlóban.

Van-e kiút?

Ezeket a problémákat fogjuk a mai esemény során megvitatni. Nagyon sikeres konferenciát kívánok mindenkinek.

John Haldane, "Towards a Philosophy of Culture"

The emergence and recent developments in the design, promulgation and rapidly increasing use of Artificial Intelligence raise a multitude of theoretical, practical and socio-cultural issues. These are already the subject of intense study across a range of scientific and human-

ities disciplines, and the focus of much public commentary and policy deliberations. Given, however, the rapid and seemingly accelerating pace and scale of current developments there is a serious question as to the feasibility of the idea that academic study can make a significant contribution to the issue of how to respond to the AI revolution. One aspect of this concerns the utility of any academic product, another is the pressure it places upon practitioners to reflect upon the nature of their disciplines and specifically the relation within them between theory and practice. Here I explore these matters as they apply to the philosophy of artificial intelligence, suggesting that this might best be conceived and practise as a form of *Kulturphilosophie*.

***

John Haldane, „A kultúra filozófiája felé”

A mesterséges intelligencia tervezésének, elterjesztésének és rohamosan növekvő használatának megjelenése és legújabb fejleményei elméleti, gyakorlati és társadalmi-kulturális kérdések sokaságát vetik fel. Ezek már most is intenzív kutatás tárgyát képezik a természettudományok és a humán tudományok különböző területein, továbbá a közéleti kommentárok és a szakpolitikai megfontolások fókuszában állnak. Mindazonáltal a jelenlegi fejlemények gyors és látszólag egyre gyorsuló üteme és léptéke komoly kérdést vet fel azzal kapcsolatban, mennyire tartható az az elképzelés, hogy az akadémiai kutatás érdemi hozzájárulást tud nyújtani ahhoz, miként kellene reagálnunk az MI-forradalomra. Ennek egyik aspektusa az akadémiai teljesítmények gyakorlati hasznosságára vonatkozik, egy másik pedig arra a nyomásra, amelyet mindez a kutatókra gyakorol: hogy reflektáljanak saját diszciplínáik természetére, és különösen arra, miként viszonyul egymáshoz bennük az elmélet és a gyakorlat. Ebben az írásban e kérdéseket vizsgálom a mesterséges intelligencia filozófiájára vonatkoztatva, és amellett érvelek, hogy ezt a területet talán leginkább a *Kulturphilosophie* egyik formájaként érdemes felfogni és művelni.

Valéria Csépe, "Cognitive Offloading: How to Preserve Human Reasoning in an AI-dominated World?"

Artificial intelligence (AI) as an integral part of our life seems to streamline daily life, including search queries, task performance, text, music, video and picture generation and even decision-making. Although most of the AI tools offer convenience and efficiency, the risk of cognitive offloading is higher than ever before. This psychological term is used for the process delegating cognitive tasks to external aids and assumed to result in diminished skills, declined critical thinking and in general as altered fundamental cognitive processes. The question often raised by researchers is weather any balanced AI usage is possible? Moreover, is the incorporation of AI essential for the advancement of our societies? Is the human intelligence at risk if AI takes it all? Is there a moderate use of AI with positive cognitive impact? If yes, how to learn it?

There are several questions and the presentation aims at answering some of them referring to reliable (means reproducible) research data when available. However, the most important issues are now just to discuss and discover a possible scenario seen from the perspective of cognitive psychology, neuroscience and educational sciences. The impact on the generation born after 2010, called alpha is known though not fully clear. However, the possible changes on those who are born after 2025, called as beta and are in cradle, can only be expected without any information on the further technological inventions coming.

The five topics in focus of the presentation are: (1) research findings on AI and critical thinking, (2) benefits and cognitive risks, (3) development of metacognitive skills to assess quality and reliability of AI-generated outputs, (4) mitigation of AI's cognitive costs; (5) implications for education.

***

Csépe Valéria, „Kognitív tehermentesítés: Hogyan őrizhető meg az emberi érvelés egy AI-dominált világban?”

Az életünk szerves részévé vált mesterséges intelligencia (AI) úgy tűnik, hogy leegyszerűsíti a mindennapokat: a keresési folyamatokat, a feladatvégzést, a szöveg-, zene-, videó- és képgenerálást, sőt még a döntéshozatalt is. Noha az AI-eszközök többsége kényelmet és hatékonyságot kínál, a kognitív tehermentesítés kockázata minden eddiginél nagyobb. Ezt a pszichológiai fogalmat arra a folyamatra használjuk, amikor kognitív feladatokat külső segédeszközökre ruházunk át, ami feltételezhetően csökkent készségekhez, hanyatló kritikai gondolkodáshoz és általában az alapvető kognitív folyamatok megváltozásához vezet. A kutatók gyakran teszik fel a kérdést: vajon lehetséges-e bármiféle kiegyensúlyozott AI-használat? Továbbá: elengedhetetlen-e az AI beépítése társadalmaink fejlődéséhez? Veszélybe kerül-e az emberi intelligencia, ha az AI mindent átvesz? Létezik-e mérsékelt, pozitív kognitív hatású AI-használat? Ha igen, hogyan sajátítható el?

Számos kérdés merül fel, és az előadás ezek közül néhányra igyekszik választ adni, ahol rendelkezésre állnak megbízható (azaz reprodukálható) kutatási adatok. A legfontosabb kérdések azonban jelenleg inkább diszkusszióra és feltárásra várnak, egy lehetséges jövőkép megrajzolására a kognitív pszichológia, az idegtudomány és az oktatás tudományainak perspektívájából. A 2010 után született generáció, az alfa esetében már vannak ismereteink, bár ezek nem teljesen egyértelműek. A 2025 után születettek – az úgynevezett béta generáció, amely most bölcsőben van – esetében azonban a lehetséges változások csak feltételezhetők, hiszen semmilyen információnk nincs arról, milyen további technológiai találmányok jelennek majd meg.

Az előadás öt kiemelt témája: (1) kutatási eredmények az AI és a kritikai gondolkodás kapcsolatáról; (2) előnyök és kognitív kockázatok; (3) metakognitív készségek fejlesztése az AI-által generált kimenetek minőségének és megbízhatóságának értékeléséhez; (4) az

AI kognitív költségeinek mérséklése; (5) következmények az oktatás számára.

Eörs Szathmáry, “How AI is Becoming More Biological”

Evolvable AI (eAI), i.e., AI systems whose components, learning rules, and deployment conditions can themselves undergo Darwinian evolution, may soon emerge from current trends in generative, agentic, and embodied AI. We argue that this possibility has been underappreciated in debates on AI safety and existential risk. Here, we ask under what technical and ecological conditions AI becomes evolvable, what kinds of behaviours are then likely to emerge, and how such systems could be governed. Drawing on biological evolution and decades of digital evolution experiments, we distinguish “breeder” scenarios, in which humans impose fitness criteria and control reproduction, from “ecosystem” scenarios, in which selection arises from open environments and control erodes. In the latter, selfish replication reliably gives rise to cheating, parasitism, deception, and manipulation, even in very simple systems.

Evolvable AI agents are expected to behave increasingly like biological organisms in cyberspace. They will not only be evolutionary units, but will metabolize, compete, cooperate, and parasitize as part of a kind of Life 2.0. I shall briefly outline how concepts such as individuality, the genotype-phenotype map, fitness, and niche construction will help analyze these concepts.

***

Szathmáry Eörs, „Hogyan válik az AI egyre inkább biológiai jellegűvé?”

Az evolúciós képességű mesterséges intelligencia (eAI) – azaz azok a mesterséges intelligencia-rendszerek, amelyek komponensei, tanulási szabályai és üzemeltetési feltételei maguk is darwini evolúción mehetnek keresztül – hamarosan megjelenhet a generatív, ügynökalapú

és megtestesült mesterséges intelligencia jelenlegi trendjeiből. Álláspontunk szerint ezt a lehetőséget eddig alulértékelték a mesterséges intelligencia biztonságával és az egzisztenciális kockázattal kapcsolatos vitákban. Jelen tanulmányban azt vizsgáljuk, hogy milyen technikai és ökológiai feltételek mellett válik a mesterséges intelligencia evolúciós képessé, milyen viselkedésformák alakulhatnak ki ilyenkor, és hogyan lehetne szabályozni ezeket a rendszereket. A biológiai evolúcióra és az évtizedek óta folyó digitális evolúciós kísérletekre támaszkodva megkülönböztetjük a „tenyésztői" forgatókönyveket – amelyekben az emberek szabnak meg alkalmassági kritériumokat és szabályozzák a szaporodást – az „ökoszisztémás" forgatókönyvektől, amelyekben a szelekció nyitott környezetekből fakad, és a kontroll gyengül. Utóbbi esetben az önző replikáció még nagyon egyszerű rendszerekben is megbízhatóan vezet csaláshoz, parazitizmushoz, megtévesztéshez és manipulációhoz.

Az evolúciós képességű mesterséges intelligencia-ügynökök várhatóan egyre inkább biológiai organizmusokhoz hasonlóan fognak viselkedni a kibertérben. Nem csupán evolúciós egységek lesznek, hanem egyfajta „Life 2.0" részeként anyagcserét folytatnak, versenyeznek, együttműködnek és parazitálnak. Röviden felvázolom, hogyan segítenek az egyediség, a genotípus–fenotípus leképezés, a rátermettség és a niche-építés fogalmai ezeknek a jelenségeknek az elemzésében.

Csaba Pléh, "Relationships between Human Thinking and Machine Based Solutions: Remarks of an Outsider"

> Outsider refers in the title to the issue that I am untouched by the recent AI craze, and avoid using it even as a tool. My remarks shall be about the past of this tradition and the generality of the moral panic around recent LLM-s.

1. *The relationship between man and technology*. The instrumental image of man. What is the specialty of the computer in this regard

and recent AI models specifically? If there is a thinking machine, maybe humans also think.
2. *What was easy and what was difficult* for computers in the past? Machines had difficulty in perception and in interpreting the social world, which in humans is based on automatized skills. Thought experiments like the Chinese Room, about the cause of the difficulty. What was the trick of these thought experiments? The reemergence of the social issue in the recent AI debates: difficulty of LLMs with relevance and empathy.
3. *Imitation of performance or process.* This is traditional issue from the levels of the Turing test, from the chess machine to artificial man (István Hernád).
4. *Present day moral panic about recent AI.* The reclaiming humanity attitude *vis a vis* Silicon Walley is similar to reclaiming humanity in the name of … religion, phenomenology, hermeneutics etc. from neural reductionism, from genetic determination, and even from classical ex-perimental psychology. What is really new in here?

***

Pléh Csaba,
„Az emberi és a gépi megoldások közötti kapcsolatok:
Egy kívülálló megjegyzései”

> A címben a ‘kívülálló’ arra utal, hogy nem érint meg a jelenlegi MI megszállottság, s arra törekszem, hogy még eszközként se használjam. Megjegyzéseim a hagyomány múltjára vonatkoznak és a korunk nagy Nyelvi Modelljeit övező erkölcsi pánik általánosabb kereteire.

1. *Emberek és technológia kapcsolata*. Az ember instrumentális képe az újkor kezdete óta velünk van. Mi ebben a számítógép s főként a mai MI modellek különlegessége? Ha vannak gondolkodó gépek, lehet, hogy az ember is gondolkodik?
2. *Mi volt régen könnyű és nehéz a számítógépek számára?* A gépeknek nehézségeket okozott az észlelés és a társas világ értelmezése,

melyek embereknél automatizált készségeken alapulnak. Gondolatkísérletek, mint például a Kínai Szoba születtek a nehézségek okainak megvilágítására. Mi is volt a gondolatkísérletek alapvető trükkje? A társas gondok újra megjelennek az MI-t körülvevő mai vitákban: a mai Nagy Nyelvi Modelleknek gondjaik vannak a relevanciával és az empátiával.

3. *A teljesítmény és a folyamat utánzása.* Hagyományos kérdés ez a különböző szintű Turing próbáktól kezdve, a sakkozógéptől a mesterséges emberig (Hernád István).

4. *A mai MI kiváltotta mai erkölcsi pánik.* Az ember mivolt visszakövetelése a Szilikon Völgytől hasonlít ahhoz, amikor az ember mivoltot visszakövetelik … a vallás, a fenomenológia, a hermenutika stb. nevében. Mentsük meg az embert a neurális redukcionizmustól, a genetikai meghatározottságtól, sőt még a klasszikus kísérleti pszichológiától is, hangzott klasszikusan. Mi is itt újdonság?

5. *A konnekcionizmus és a gépi tanulás kihívása.* Fél évszázada velünk él egy statisztikai versus szabálykövető hozzáállás például a nyelv elemzésében. A szabály-elv a korai 1960-as években könnyű győzelmet aratott, mígnem az 1980-as évek közepén a jobb gépek és a mára Nobel díjjal jutalmazott kapcsolatelvű statisztikus programok révén a statisztikai hozzáállás diadalmasan visszatért. Sokan vagyunk, Steven Pinker indítását követve, például Gary Marcus, akik úgy gondoljuk, hogy a kettősség magára az emberi gondolkodásra jellemző. Mi emberek egyszerre Boltzmann gépek és Frege gépek vagyunk. Vajon lesznek-e olyan gépi megoldások, melyekbe beilleszkedik maga ez a kettősség? Lesz-e algebrai alapú MI?

András Falus,
“Philosophical-Psychological Perspectives on Medical AI”

Artificial intelligence (AI) has transformative potential across diagnostics, therapeutics, drug development, and biomarker research. In medical diagnostics, AI enables early detection and risk assessment through the integration of multimodal data. In therapeutics, it sup-

ports the selection of treatments, dosing, and monitoring with adaptive, data-driven decision-making. In drug and vaccine development, it accelerates lead discovery, de novo design, and safety evaluations, thereby shortening development timelines.

Beyond efficiency, we should expect AI to balance human values – autonomy, dignity, and trust – with value-aligned design and clear accountability. From a philosophical perspective, the horizons of research and clinical practice include, among other things, the frontiers of human knowledge; the socio-psychological dimension of doctor–patient relationships should be treated as educational priorities. Real benefits arise when AI complements clinical decision-making while properly addressing generalizability, data privacy, and equitable access.

***

Falus András,
„Filozófiai-pszichológiai nézőpontok az orvosi AI-hoz”

A mesterséges intelligencia (AI) átalakító potenciállal bír a diagnosztika, terápiák, gyógyszerfejlesztés és biomarker-kutatás terén. Az orvosi diagnosztikában az AI korai felismerést és kockázatbecslést tesz lehetővé multimodális adatok integrációjával. A terápiákban támogatja a kezelések kiválasztását, az adagolást és a monitorozást adaptív, adat-alapú döntéshozatallal. A gyógyszer- és vakcinafejlesztésben felgyorsítja a lead-kutatást, a de novo tervezést és a biztonsági értékeléseket, így rövidíti a fejlesztési időt.

Az AI-tól a hatékonyság mellett az emberi értékek – autonómia, méltóság és bizalom – egyensúlyát kell elvárjuk, értékekhez igazított tervezéssel és egyértelmű felelősség-meghatározással. A kutatási és gyógyítási távlatok filozófiai vetületben többek között az emberi megismerés távlatait, a szociálpszichológiai dimenzió a doktor–beteg kapcsolatokat kell, hogy edukálható prioritásként kezelje. A valódi előnyök akkor jelentkeznek, ha az AI kiegészíti a klinikai döntéshozatalt, megfelelően kezelve az általánosíthatóságot, adatvédelmet és a hozzáférés egyenlőségét.

Xu Wen, "Between Dao and Machine:
Towards a Chinese Philosophy of Artificial Intelligence –
Rethinking Agency, Harmony, and Human Responsibility
in the Age of AI"

This talk develops a Chinese philosophical approach to artificial intelligence by placing machine intelligence between Dao (a key concept in Chinese Philosophy) and human responsibility. Rather than treating AI only as a technical instrument or an existential threat, it asks how AI transforms the relation among human agency, cosmic order, social harmony, and ethical cultivation. Drawing on Daoist reflections on spontaneity, Confucian ideas of self-cultivation and relational responsibility, as well as the classical notion of harmony without uniformity, the talk argues that AI should not be understood as a rival subject replacing humanity, but as a powerful relational force that reorganizes human action, knowledge, and social life. The central question is therefore not whether machines can think like humans, but how humans should live, judge, and act with intelligent machines. A Chinese philosophy of AI calls for technological development guided by balance, responsibility, humility, and care for the shared future of humanity in an increasingly automated world today.

***

Xu Wen, „A Dao és a gép között:
Egy kínai mesterséges-intelligencia filozófia felé –
A cselekvés, harmónia és emberi felelősség újragondolása
az MI korában"

Ez az előadás a mesterséges intelligenciára vonatkozó kínai filozófiai megközelítést dolgoz ki azáltal, hogy a gépi intelligenciát a Dao (a kínai filozófia egyik alapfogalma) és az emberi felelősség közé helyezi. Ahelyett, hogy az MI-t pusztán technikai eszközként vagy egzisztenciális fenyegetésként kezelnénk, azt vizsgálja, miként ala-

kítja át az MI az emberi cselekvőképesség, a kozmikus rend, a társadalmi harmónia és az etikai önművelés viszonyát.

A daoista spontaneitásról szóló reflexiókra, a konfuciánus önkultiváció és relációs felelősség fogalmaira, valamint a klasszikus „egység nélküli harmónia" eszméjére támaszkodva az előadás amellett érvel, hogy az MI-t nem szabad az emberiséget helyettesítő rivális szubjektumként felfogni, hanem olyan erőteljes relációs tényezőként, amely újraszervezi az emberi cselekvést, tudást és társadalmi életet.

A központi kérdés ezért nem az, hogy a gépek képesek-e úgy gondolkodni, mint az emberek, hanem az, hogy az emberek miként éljenek, ítéljenek és cselekedjenek az intelligens gépekkel együtt. A mesterséges intelligencia kínai filozófiája olyan technológiai fejlődést sürget, amelyet az egyensúly, a felelősség, az alázat és az emberiség közös jövőjéért viselt gond határoz meg a mindinkább automatizált világban.

James E. Katz,
"Saying Thanks to No One:
Buber, Kant, and the Case for Apparatgeist"

Martin Buber's distinction between I-Thou and I-It relations remains a standard resource for theorizing how we address other beings: I-It denotes instrumental, functional engagement with objects handled and set aside; I-Thou denotes mutual, presential relation, irreducible to use. Kant, working within a different framework, denies that non-rational beings can be objects of direct duty. He nonetheless holds that we have indirect duties regarding them: cruelty toward animals wrongs no one but the agent, hardening dispositions that later surface in our treatment of other persons.

Both frameworks are increasingly invoked in discussions of human interaction with artificial intelligence (AI), yet neither is sufficient on its own. Many users address generative AI systems with courtesy, apologizing or thanking, while explicitly disavowing any

belief that these systems have interests or the capacity to be addressed. Buber's framework struggles to place this behavior: no genuine mutuality is claimed, so it is not I-Thou, yet it exceeds what pure I-It instrumentality would predict, since strict efficiency would eliminate such courtesies rather than preserve them. Kant's account of indirect duty offers a partial diagnosis: it relocates the obligation from the object to the agent. But the account applies equally to any non-rational object, so it cannot explain why AI systems, and not other non-sentient artifacts, have become common sites for this behavior.

I argue instead, drawing on my own Apparatgeist framework and the related concept of perpetual contact (developed with M. Aakhus), that what marks certain artifacts out for this kind of address is not machine cognition but ongoing, one-sided availability: the fact that the artifact is always on hand for interaction regardless of the human's mood, timing, or need. Contact of this kind, sustained over time, explains why some artifacts draw this kind of address and others do not. I close by weighing this account, in which politeness protects the agent's existing character, against a rival reading drawn from Ubuntu personhood, in which such address instead builds character through the exchange itself. Evidence about when and why users extend or withhold courtesy toward AI systems could help decide which account better describes what is actually happening, even though it cannot by itself settle which account of duty, or of personhood, is normatively correct. Distinguishing those two questions is itself unfinished business for a philosophy of AI.

***

James E. Katz, „Köszönet senkinek: Buber, Kant és indok az Apparatgeist mellett"

Martin Buber Én–Te és Én–Az viszonyok közti megkülönböztetése továbbra is alapvető erőforrás annak elméletében, miként fordulunk más létezők felé: az Én–Az az instrumentális, funkcionális viszonyt jelöli, amelyben tárgyakat kezelünk és félreteszünk; az Én–Te pedig a kölcsönös, jelenléti kapcsolatot, amely nem redukálható puszta hasz-

nálatra. Kant – egészen más keretben – tagadja, hogy nem-racionális lények iránt közvetlen kötelességeink lennének. Mindazonáltal úgy véli, hogy közvetett kötelességeink vannak velük kapcsolatban: az állatokkal szembeni kegyetlenség senkit sem sért, csak magát a cselekvőt, mivel olyan hajlamokat keményít meg, amelyek később más személyekkel való bánásmódunkban törnek felszínre.

Mindkét keretet egyre gyakrabban hívják segítségül az ember és a mesterséges intelligencia (MI) közti interakciók tárgyalásában, ám egyik sem elégséges önmagában. Sok felhasználó udvariasan szólítja meg a generatív MI-rendszereket, bocsánatot kér vagy köszönetet mond, miközben kifejezetten tagadja, hogy e rendszereknek érdekeik vagy megszólíthatósági képességük volna. Buber kerete nehezen talál helyet ennek a viselkedésnek: valódi kölcsönösséget nem állítanak, tehát nem Én–Te, ugyanakkor meghaladja azt, amit a tiszta Én–Az instrumentális viszony előre jelezne, hiszen a szigorú hatékonyság kiküszöbölné az efféle udvariassági gesztusokat, nem pedig megőrizné őket. Kant közvetett kötelességről szóló elmélete részleges diagnózist kínál: a kötelezettséget az objektumról a cselekvőre helyezi át. Csakhogy az elmélet minden nem-racionális tárgyra egyformán vonatkozik, így nem magyarázza meg, miért éppen az MI-rendszerek – és nem más, szintén nem érző artefaktumok – váltak e viselkedés gyakori helyszíneivé.

Én ehelyett – saját Apparatgeist elméletemet és a hozzá kapcsolódó, M. Aakhusszal közösen kidolgozott *perpetual contact* (állandó kontaktus) fogalmát felhasználva – amellett érvelek, hogy azokat az artefaktumokat, amelyek felé az efféle megszólítás irányul, nem a gépi kogníció különíti el, hanem az egyoldalú, folyamatos rendelkezésre állás: az a tény, hogy az artefaktum mindig készen áll az interakcióra, függetlenül az ember hangulatától, időzítésétől vagy szükségletétől. Az ilyen jellegű, időben tartós kontaktus magyarázza, miért váltanak ki egyes artefaktumok ilyen megszólítást, míg mások nem.

A tanulmány végén mérlegelem ezt az értelmezést – amely szerint az udvariasság a cselekvő meglévő jellemét védi – egy riválisolvasattal szemben, amely az Ubuntu-személyiség fogalmából indul

ki, és amely szerint az efféle megszólítás éppen a cserefolyamat révén építi a jellemet. Az hogy a felhasználók mikor és miért tanúsítanak vagy tartanak vissza udvariasságot az MI-rendszerekkel szemben, segíthet eldönteni, melyik értelmezés írja le pontosabban, mi is történik valójában – még ha önmagában nem is képes eldönteni, mely kötelesség- vagy személyiségfogalom normatíve helyes. E két kérdés megkülönböztetése maga is befejezetlen feladat az MI filozófiája számára.

Attila Orbán, “Memory, Tradition and Artificial Intelligence: Towards an Ontology of Cultural Continuity”

Contemporary discussions on artificial intelligence often focus on technological capabilities, ethical risks, or future social transformations. This paper approaches AI from a different perspective by examining it within the long history of human tradition and cultural memory.

The central thesis argues that tradition should not be understood merely as the preservation of customs or inherited knowledge. Rather, tradition represents a fundamental mode through which ordered structures preserve themselves through time. Biological inheritance, oral cultures, writing, canonical texts, printing, electronic media, and artificial intelligence can all be interpreted as successive forms of memory that enable the continuity of human communities.

The paper begins with an ontological reflection on life itself. Contemporary biology demonstrates that life is distinguished from non-life not by the chemical elements composing it, but by their highly organized dynamic structure. This insight provides a philosophical analogy for understanding tradition: just as living organisms preserve biological order through genetic memory, human cultures preserve symbolic order through cultural memory.

The study then follows the historical transformation of memory from myth and ritual to literacy and digital technologies. Myths

are interpreted not as primitive explanations of nature but as early structures of collective memory. The invention of writing externalized memory; canonization stabilized it; printing democratized it; electronic media accelerated its circulation; and artificial intelligence introduces a new stage in which cultural memory becomes computationally accessible.

The presentation does not claim that artificial intelligence replaces human tradition. Instead, it argues that AI constitutes a new technological medium within the historical evolution of memory while leaving unresolved the uniquely human dimensions of meaning, responsibility, and moral judgment.

***

Orbán Attila,

„Emlékezet, hagyomány és mesterséges intelligencia: a kulturális folytonosság ontológiai vizsgálata”

A mesterséges intelligenciáról szóló kortárs diskurzusok gyakran a technológiai képességekre, az etikai kockázatokra vagy a jövőbeli társadalmi átalakulásokra összpontosítanak. Jelen előadás más nézőpontból közelít a mesterséges intelligenciához: az emberi hagyomány és a kulturális emlékezet hosszú történetének összefüggésében vizsgálja.

Az előadás központi tézise szerint a hagyomány nem érthető meg pusztán a szokások vagy az örökölt tudás megőrzéseként. A hagyomány sokkal inkább olyan alapvető létmód, melynek révén a rendezett struktúrák fennmaradnak az időben. A biológiai öröklődés, a szóbeli kultúrák, az írás, a kanonikus szövegek, a könyvnyomtatás, az elektronikus média és a mesterséges intelligencia egyaránt az emlékezet egymást követő formáiként értelmezhetők, amelyek lehetővé teszik az emberi közösségek folytonosságát.

Az előadás magára az életre irányuló ontológiai reflexióból indul ki. A kortárs biológia arra mutat rá, hogy az élőt az élettelentől nem az őket alkotó kémiai elemek, hanem ezen elemek magas fokon szervezett, dinamikus struktúrája különbözteti meg. Ez a felismerés

filozófiai analógiát kínál a hagyomány megértéséhez. Ahogyan az élő szervezetek a genetikai emlékezet révén őrzik meg a biológiai rendet, úgy az emberi kultúrák a kulturális emlékezet révén tartják fenn a szimbolikus rendet.

A vizsgálat ezt követően végigkíséri az emlékezet történeti átalakulását a mítoszoktól és a rítusoktól az írásbeliségen át a digitális technológiákig. A mítoszokat nem a természet kezdetleges magyarázataiként, hanem a kollektív emlékezet korai struktúráiként értelmezi. Az írás feltalálása külső hordozóra helyezte az emlékezetet; a kanonizáció megszilárdította, a könyvnyomtatás széles körben hozzáférhetővé tette, az elektronikus média felgyorsította áramlását, a mesterséges intelligencia pedig egy új szakaszt nyit, amelyben a kulturális emlékezet számítógépes úton válik hozzáférhetővé és értelmezhetővé.

Az előadás nem azt állítja, hogy a mesterséges intelligencia felváltja az emberi hagyományt. Érvelése szerint inkább olyan új technológiai közeget alkot, amely az emlékezet történeti fejlődésébe illeszkedik, miközben továbbra is megoldatlanul hagyja a jelentés, a felelősség és az erkölcsi ítéletalkotás sajátosan emberi dimenzióit.

A mesterséges intelligencia filozófiáját ezért nem csupán az intelligens gépek filozófiájaként kell felfognunk, hanem azon változó formák filozófiájaként is, amelyek révén az emberiség megőrzi, újrarendezi és továbbadja kollektív emlékezetét.

Gabriella Manzenreiter, “From Creation to Co-Creation: Philosophical Perspectives on Human and Machine Creativity”

Is AI a genius, or is it rather a creative agent? The British artist Harold Cohen thought that a computer program might represent knowledge that leads to the act of making art. In 1973, he began developing AARON, a computer program designed to autonomously

produce paintings and drawings. This project represents an early experiment in computational creativity.

The emergence of such an autonomous generative system raised fundamental questions about the nature of art. The creative process resulted in tangible outcomes, suggesting that machines are capable of generating and representing in ways that appear analogous to artistic practice. But what do these creations represent? Furthermore, what constitutes creativity and artistic creation?

In this paper, I explore the dimensions of human and machine creativity and the relationship between them, with the aim of creating a philosophical and conceptual framework that highlights their similarities, differences, and possible intersections, as well as the socio-cultural implications of machine creativity in the age of human-machine co-creativity.

***

Manzenreiter Gabriella,
„Az alkotástól a kokreációig:
az emberi és a gépi kreativitás filozófiai perspektívái”

A mesterséges intelligencia zseni, vagy inkább kreatív ágens? Harold Cohen brit művész úgy vélte, hogy egy számítógépes program olyan tudást reprezentálhat, amely elvezet a művészi alkotás aktusához. Cohen 1973-ban kezdte el fejleszteni az AARON nevű számítógépes programot, amelyet festmények és rajzok autonóm létrehozására tervezett. Ez a projekt a számítógépes kreativitás egyik korai kísérletét jelenti.

Egy ilyen autonóm generatív rendszer megjelenése alapvető kérdéseket vetett fel a művészet természetével kapcsolatban. A kreatív folyamat kézzelfogható eredményeket hozott létre, ami arra utalt, hogy a gépek képesek olyan módon generálni és reprezentálni, amely analógnak tűnik a művészeti gyakorlattal. Fel kell tennünk azonban a kérdést: mit reprezentálnak ezek az alkotások? Továbbá mi határozza meg a kreativitást és a művészi alkotást?

Előadásomban az emberi és a gépi kreativitás dimenzióit, valamint ezek egymáshoz való viszonyát vizsgálom. Célom egy olyan filozófiai és fogalmi keret kialakítása, amely kiemeli hasonlóságaikat, különbségeiket és lehetséges metszéspontjaikat, továbbá vizsgálja a gépi kreativitás társadalmi és kulturális vonatkozásait az ember–gép kokreativitás korszakában.

George Pór, "Can AI Herald the Good Society?"

Public and academic debate on artificial intelligence is increasingly polarized between AI-sceptical and AI-utopian perspectives. The risks associated with uncontrolled AI are very real under present socio-economic conditions shaped by competition among corporations and nation-states. At the same time, it is equally plausible that, under the right conditions, AI could become an enabling force for a new era of abundance, human flourishing, and the democratization of knowledge. Beneath these competing positions lie fundamental philosophical questions concerning what it means to be human, what constitutes a good society, and how collaborative hybrid intelligence (CHI) should be understood. This paper explores these conceptual and philosophical foundations as a basis for examining the conditions under which AI might contribute to desirable societal futures.

***

George Pór, „Hírnöke lehet-e az MI a jó társadalomnak?"

A mesterséges intelligenciáról folyó nyilvános és tudományos vita egyre inkább polarizálódik az MI-szkeptikus és MI-utópikus nézőpontok között. Az ellenőrizetlen MI-hez kapcsolódó kockázatok nagyon is valósak a vállalatok és nemzetállamok közötti versengés által formált jelenlegi társadalmi-gazdasági viszonyok között. Ugyanakkor az is reális lehetőség, hogy megfelelő feltételek mellett az MI a bőség, az emberi kiteljesedés és a tudás demokratizálódásának új kor-

szakát előmozdító erővé váljon. E versengő álláspontok mélyén alapvető filozófiai kérdések húzódnak meg: mit jelent embernek lenni, mi tekinthető jó társadalomnak, és hogyan értelmezhető a kollaboratív hibrid intelligencia (Collaborative Hybrid Intelligence, CHI). Az előadás ezeket a fogalmi és filozófiai alapokat vizsgálja annak feltárására, hogy milyen feltételek mellett járulhat hozzá az MI kívánatos társadalmi jövők megteremtéséhez.

Neville Chi Hang Li, "The New Tower of Babel: The Modern Attempt to Reach unto Heaven through the AI Boom"

Pope Leo XIV's first encyclical describes humanity at a crossroads regarding the rapid development of artificial intelligence (AI). The letter warns us about the "Babel syndrome" in the recent AI boom that idolizes profit maximization by sacrificing human dignity for efficiency.

The paper debunks the myth of cost reduction by AI on two fronts: 1) the artificial inflation of AI usage costs more economically and environmentally and 2) the social values of human-oriented jobs out-weigh the obsession with efficiency. It further argues that the rising wages of these high-touch sectors, e.g. healthcare and education, uphold the diverse human expressions of our morality and relationships, laying a solid foundation for the governance of AI for the common good of all people.

***

Neville Chi Hang Li, „Az új Bábel tornya: A modern kísérlet, hogy az AI-boom révén az égig érjünk"

Leo XIV. pápa első enciklikája szerint az emberiség válaszúthoz érkezett a mesterséges intelligencia (MI) gyors fejlődésével kapcsolatban. A levél figyelmeztet a „Bábel-szindrómára" az újkori MI-láz-

ban, amely a profitmaximalizálást bálványozza, s közben az emberi méltóságot áldozza fel a hatékonyság oltárán.

Az előadás két fronton is leleplezi az MI-vel kapcsolatos költségcsökkentés mítoszát: 1) az MI-használat mesterséges felduzzasztása gazdaságilag és környezetileg is többe kerül, és 2) az emberközpontú munkák társadalmi értéke felülmúlja a hatékonyságmániát. Továbbá amellett érvelek, hogy az ún. „high-touch" szektorok – például az egészségügy és az oktatás – növekvő bérei fenntartják erkölcsünk és kapcsolataink sokféle emberi kifejeződését, s ez szilárd alapot teremt az MI közjó érdekében történő szabályozásához.

Balázs Szegedy, "On the Mathematics of AI and Its Philosophical Implications"

Deep learning has become the most successful paradigm in artificial intelligence and the principal mathematical and technological engine behind the recent AI revolution. Its central idea is learning: rather than explicitly specifying the rules by which an intelligent system should operate, we construct highly expressive mathematical models and allow their internal structure to emerge through optimization over data.

This paradigm marks a significant departure from classical approaches to intelligence. A learned system is not assembled from a transparent collection of human-designed rules; its behavior arises from the interaction of vast numbers of parameters distributed across a high-dimensional space. This helps explain both the remarkable flexibility of modern AI and its characteristic "black-box" nature: the system can exhibit sophisticated behavior even when no simple symbolic account of its internal computation is available.

This shift may also have deeper philosophical significance. Intelligence has traditionally been closely associated with logic, symbolic manipulation, and explicit reasoning. The success of contemporary AI suggests that these may not constitute the deepest mathematical layer of intelligence. Logical and symbolic reasoning may

instead be emergent phenomena arising from a richer underlying substrate – continuous rather than discrete, distributed rather than localized, and probabilistic rather than strictly rule-governed.

From this perspective, the mathematics of deep learning may offer more than a method for engineering artificial systems. It may provide clues about the mathematical organization of natural intelligence itself. The emerging picture invites us to reconsider the relationship between learning, representation, reasoning, and logic, and raises the possibility that what we recognize as rational thought is a structured, low-dimensional manifestation of a much richer mathematical space beneath it.

***

Szegedy Balázs, „Az MI matematikája és filozófiai következményei”

A mélytanulás napjainkra a mesterséges intelligencia legsikeresebb paradigmájává vált, és egyben a közelmúlt AI-forradalmának fő matematikai és technológiai motorjává. Központi gondolata a tanulás: ahelyett, hogy kifejezetten megadnánk azokat a szabályokat, amelyek szerint egy intelligens rendszernek működnie kell, rendkívül kifejező matematikai modelleket konstruálunk, és belső struktúrájukat az adatokon végzett optimalizáció révén hagyjuk kialakulni.

Ez a paradigma jelentős eltérést mutat az intelligencia klasszikus megközelítéseihez képest. A tanult rendszer nem egy átlátható, ember által tervezett szabályokból összeállított konstrukció; viselkedése hatalmas számú paraméter kölcsönhatásából ered, amelyek egy magas dimenziószámú térben oszlanak el. Ez magyarázza a modern MI rendkívüli rugalmasságát, valamint jellegzetes „fekete doboz” természetét: a rendszer akkor is képes kifinomult viselkedést mutatni, amikor belső számításának nincs egyszerű, szimbolikus leírása.

Ez az elmozdulás mélyebb filozófiai jelentőséggel is bírhat. Az intelligenciát hagyományosan szorosan kötöttük a logikához, a szimbolikus manipulációhoz és az explicit érveléshez. A kortárs MI

sikere azt sugallja, hogy ezek talán nem az intelligencia legmélyebb matematikai rétegét alkotják. A logikai és szimbolikus következtetés inkább olyan emergens jelenség lehet, amely egy gazdagabb, alapvetőbb alapszubsztrátumból emelkedik ki – amely folytonos a diszkrét-tel szemben, elosztott a lokalizálttal szemben, és valószínűségi a szigorúan szabályvezérelt helyett.

Ebből a perspektívából a mélytanulás matematikája többet kínálhat, mint mesterséges rendszerek mérnöki megalkotásának módszerét. Utalhat a természetes intelligencia matematikai szerveződésére is. A kialakuló kép arra ösztönöz, hogy újragondoljuk a tanulás, a reprezentáció, a következtetés és a logika viszonyát, és felvessük annak lehetőségét, hogy amit racionális gondolkodásként ismerünk, az egy sokkal gazdagabb matematikai tér strukturált, alacsony dimenziójú megnyilvánulása.

Notes on contributors – az előadások szerzői

Valéria Csépe is an internationally renowned psychologist and neuroscientist, with extensive experience in research (language, music, classical and digital literacy, typical and atypical cognitive development), higher education, and science policy. She is full member of the Hungarian Academy of Sciences, member of the Academia Europaea. She served as deputy secretary general of the Academy (2008–2014, the first female of the elected CEOs) and as president of the Hungarian Accreditation Committee (MAB) for two terms (2016–2025), during which she played a key role in strengthening the quality assurance system of the Hungarian higher education and its international engagement.

Valéria Csépe is a research professor emeritus of the RCNS Brain Imaging Centre, professor emeritus of the University of Pannonia. She is a highly cited researcher, mentor and supervisor of many early-stage researchers. She has significant international experience, both through her scientific work and her active involvement in European and global professional communities. Her career bridges academic excellence and strategic leadership with a strong commitment to advancing quality, research, and innovation.

***

Csépe Valéria (1951) magyar pszichológus, egyetemi tanár, a Magyar Tudományos Akadémia rendes tagja. Az elemi akusztikus információfeldolgozás, valamint a kognitív pszichológia és idegtudomány neves kutatója. 2008 és 2014 között az MTA főtitkárhelyettese. 2014-

től az MTA TTK Agyi Képalkotó Központ kutatóprofesszora, a BME TTK, valamint a Pannon Egyetem egyetemi tanára Kutatói és oktatói munkája mellett 1996 és 2000 között a Magyar Pszichológiai Társaság főtitkárhelyettesi pozícióját töltötte be. 1996-tól az Audiology Neuro-Otology című tudományos folyóirat szerkesztő bizottságának tagja, illetve 2002 és 2008 között a Journal of Applied Psycholinguistics szerkesztője volt. 2010 és 2016 között az EU DG Research and Innovation mellett működő Helsinki Group on Gender in Research and Innovation magyar delegáltja. 2012 és 2018 között az ICSU (Nemzetközi Tudományos Tanács) stratégiai bizottságának tagja.

András Falus is a professor emeritus and a full member of the Hungarian Academy of Sciences as well as that of the Academia Europaea. He is a retired researcher and educator affiliated with Semmelweis University and Pázmány Péter Catholic University. He has mentored more than 30 PhD students. With expertise in molecular genetics and cell biology, he studies and develops AI-based solutions, including vaccine development, to advance biomedicine. He has a particular interest in the philosophical and ethical aspects of AI in medical research.

***

Falus András professzor emeritus, a Magyar Tudományos Akadémia és az Academia Europaea rendes tagja. Nyugalmazott kutató és oktató, a Semmelweis Egyetemhez és a Pázmány Péter Katolikus Egyetemhez kötődik. Több mint 30 doktori (PhD) hallgató munkáját irányította. Molekuláris genetikai és sejtbiológiai szaktudására támaszkodva mesterséges intelligencia-alapú megoldásokat – többek

között vakcinafejlesztési alkalmazásokat – dolgoz ki az orvostudomány fejlődésének előmozdítására. Különös érdeklődéssel fordul az orvosi kutatásban alkalmazott mesterséges intelligencia filozófiai és etikai kérdései felé.

John Joseph Haldane KCHS FRSE FRSA (born 19 February 1954) is a Scottish philosopher, commentator and broadcaster. He is a former papal adviser to the Vatican. He is credited with coining the term "analytical Thomism" and is himself a Thomist in the analytic tradition.

Haldane is associated with The Veritas Forum and is the current chair of the Royal Institute of Philosophy. He is Vice President of the Catholic Union of Great Britain.

James E. Katz, PhD., Dr. h.c., is the Feld Professor of Emerging Media Studies at Boston University. His pioneering publications on artificial intelligence (AI), social media, mobile communication, and robot–human interaction have been internationally recognized and translated into a dozen languages. His most recent book, co-edited with Katie Schiepers and Juliet Floyd, is *Nudging Choices Through Media: Ethical and Philosophical Implications for Humanity* (Palgrave Macmillan). He holds two patents and is the winner of numerous awards including the prestigious Frederick Williams Prize for Contributions to the Study of Communication Technology, given by the International Communication Association. According to Google Scholar, he has been cited more than 18,000 times. E-mail: <profjameskatz@gmail.com>.

Neville Li, Ph.D., is Program Coordinator of the International Studies at Saint Louis University Madrid. His PhD research examines the securitization of population dynamics, and his current research interests focus on politics in the Asia-Pacific, political communication, and political economy.
E-mail: neville.li@slu.edu.

Gabriella Manzenreiter (1993) is a curator and Ph.D. student at the Doctoral School of Philosophy at the University of Pécs, Hungary. Her research areas include the philosophy of art, art theory, and aesthetics. Her primary research focuses on theoretical questions regarding the application of artificial intelligence in creative processes.

Research:
https://medbiotech.academia.edu/GabriellaManzenreiter
E-mail: manzenreiter.gabriella@edu.pte.hu.

***

Manzenreiter Gabriella (1993) kurátor és a Pécsi Tudományegyetem Filozófia Doktori Iskolájának doktorandusza. Kutatási területei a művészetfilozófia, a művészetelmélet és az esztétika. Fő kutatási tevékenysége a mesterséges intelligencia kreatív folyamatokban történő alkalmazásával kapcsolatos művészetelméleti és esztétikai kérdések vizsgálatára terjed ki.
Kutatás: https://medbiotech.academia.edu/GabriellaManzenreiter
E-mail: manzenreiter.gabriella@edu.pte.hu.

Kristóf [J. C.] Nyíri, born 1944, is Full Member of the Hungarian Academy of Sciences. He has held professorships at various universities in Hungary and abroad. He was Leibniz Professor of the University of Leipzig for the winter term 2006/07. His main fields of research are the philosophy of conservatism, the history of philosophy in the 19th and 20th centuries, the impact of communication technologies on the organization of ideas and on society, the philosophy of images, and the philosophy of time.
For further information see:
www.hunfi.hu/nyiri and https://mta.acade mia.edu/KristofNyiri.
E-mail: nyirik@gmail.com.

***

Nyíri Kristóf (1944) a Magyar Tudományos Akadémia rendes tagja. Professzori állásokat töltött be különböző magyarországi és számos külföldi egyetemen. A Lipcsei Egyetem Leibniz-professzora volt a 2006/07-es téli szemeszterben. Fő kutatási területei: a konzervatizmus filozófiája, a 19. és 20. század filozófiatörténete, a kommunikációs technológiák hatása az eszmék és a társadalom szerveződésére, a képek filozófiája és az idő filozófiája. További információk:
www.hunfi.hu/nyiri és https://mta.academia.edu/KristofNyiri.
E-mail: nyirik@gmail.com.

Attila Orbán is a Teaching Assistant at the Department of Cultural Theory and Applied Communication Studies, University of Pécs, Hungary, and a PhD candidate at the Doctoral School of Philosophy of the same university. His research focuses on the social-philosophical implications of digital culture, with particular emphasis on the transformation of the concepts of community and tradition in the

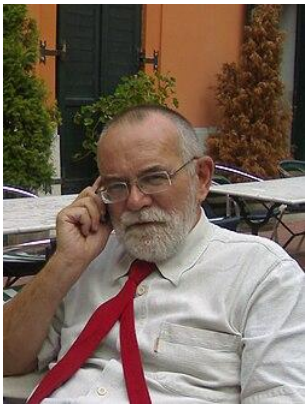
context of technological change, mediatization, and digital communication. His doctoral dissertation, *Changing Concepts of Community and Tradition in Digital Culture*, investigates how emerging communication technologies reshape collective identity, cultural memory, social relationships, and the transmission of traditions.

Orbán's academic background includes studies in theology, psychology, communication, political science, leadership and organizational studies, development policy, and philosophy. His work explores the relationship between communication technologies and human thought, the transition from oral and literate cultures to digital environments, and the evolving forms of collective memory and cultural continuity. His research draws upon the works of Walter J. Ong, Harold Innis, Marshall McLuhan, Jan Assmann, Hans-Georg Gadamer, and Kristóf Nyíri.

Alongside his academic activities, he has extensive experience in community leadership, education, and public service. He currently serves as President of the General Assembly of Tolna County, Hungary. – Research interests: digital culture, philosophy of communication, media ecology, community theory, cultural memory, tradition studies, technological determinism, artificial intelligence and society. E-mail: orbanattilapaks@gmail.com.

***

Orbán Attila (1973) teológus, a Pécsi Tudományegyetem Filozófia Doktori Iskolájának doktorandusza, valamint a PTE Kultúratudományi, Pedagógusképző és Vidékfejlesztési Karának egyetemi tanársegédje. Tanulmányokat folytatott a teológia, a pedagógia, a pszichológia, a politikai kommunikáció, a vezetés- és szervezéstudomány, valamint a fejlesztéspolitika területén. 2007-2008-ban dél-afrikai zulu közösségek körében tanulmányozta a hagyományos kultúra és gondolkodás sajátosságait. Kutatási területe a hagyomány ontológiai megnyilvánulásainak vizsgálata a kommunikációs formák történeti átalakulásának összefüggésében. Kutatómunkája részeként részt vett Nyíri Kristóf *Kép és idő* című szemináriumán.
E-mail: orbanattilapaks@gmail.com.



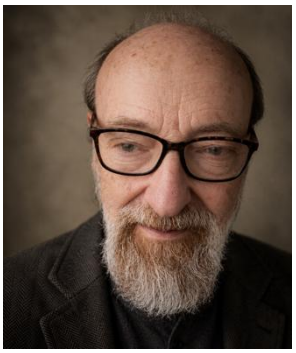
Csaba Pléh (born 1945) is a Hungarian psychologist and linguist. He is a member of the Hungarian Academy of Sciences, and of the Academia Europeae (London). His empirical research mainly concentrated on the cross-linguistic study of language processing and language development, and issues of developmental language disorders. His theoretical work relates to history of psychology and cognitive science, including the fate of machine based models of human thought. He has been working at CMU, UCSD, Vienna University, Stanford, Indiana U., Rutgers, Lyon, Eszterházy College, and Eötvös U and Technical U of Budapest. Recently, he is professor emeritus at CogSci Central Europan University, Budapest.
E-mail: Vispleh@ceu.edu.

Conference organizers' remark: Prof. Pléh introduced a talk on AI given by Gary Marcus at the Hungarian Academy of Sciences on May 10, 2024. For the introduction and the talk see:

https://share.google/h8tkBigUn9gpSyYNJ.

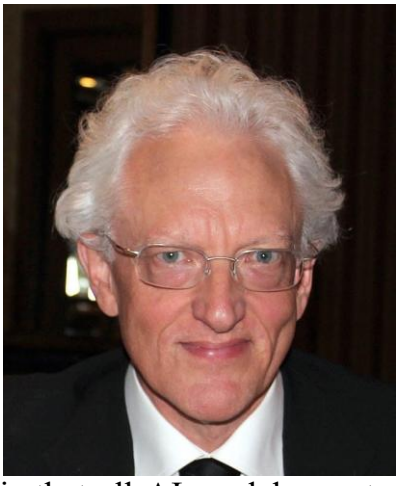
***

Pléh Csaba (sz.1945) magyar pszichológus és nyelvész. A Magyar Tudományos Akadémia és az Academia Europeae (London) tagja. Empirikus kutatásai az emberi nyelvfeldolgozás és nyelvelsajátítás nyelvközi összehasonlítására valamint a nyelvi fejlődés zavaraira összpontosítottak. Elméleti munkái a pszichológia és a kognitív tudomány történetével foglalkoznak, beleértve az emberi gondolkodás gépi alapú értelmezéseink sorsát. Az ELTE, ME, SzTE, BME és az Eszterházi Károly Főiskola mellett tanított illetve kutatott a CMU, UCSD, Bécsi, Stanford, Indiana, Rutgers és Lyoni Egyetemeken. Jelenleg a Central Europan University, Budapest Kognitív Tudományi Tanszékének nyugalmazott professzora.
E-mail: Vispleh@ceu.edu.



George Pór is the Founder and Director of Research at Future HOW—Center of Action Research for Regenerative Futures, where he developed Generative Action Research methodology. A metamodern political sociologist and resarcher in the collaborative hybrid intelligence of symbiotically linked human and wisdom-fostering AI agents, he is also a Distinguished Fellow at Schumacher Institute. His transdisciplinary reasearch spans the AI-augmented wisdom-praxis, transformative communities of practice, and strengthening resilience from personal to planetary scales. He held appointments at several European and US universities, co-authored 10 books and is working on two forthcoming books: (a) The Rise of Compassionate AI and (b) AI for a DELIBERATELY DEVELOPMENTAL SOCIETY: Exploring the Emancipatory Potential of Artificial Intelligence. E-mail: george.por@gmail.com.

Barry Smith is one of the most widely cited contemporary philosophers. He has made influential contributions to ontology, and especially to applications of ontology in the biomedical, industrial and security domains. He edited parts 1 and 2 of the ISO/IEC 21838 ontology standard, which is the first example of a piece of philosophy subjected to the ISO standardization process. He also co-authored a book entitled *Why Machines Will Never Rule The World*, which attempts to dismantle the current hype on behalf of so-called "artificial intelligence". Its principal thesis is that all AI models must execute on computers, which means that they must rely on mathematical algorithms which are Turing computable. (All computers are Turing machines.) This places severe limits on what AI models can achieve, for example when it comes to



the emulation of the behaviour of that complex system which is the human brain.
E-mail: ifomis@gmail.com.

Eörs Szathmáry (born 1959) is a Hungarian theoretical evolutionary biologist, widely known for his foundational work together with the late John Maynard Smith on the major evolutionary transitions. He has worked on the origin of life, systems chemistry, replicator theory, the origin of the genetic code, the emergence of the eukaryotic cell, the origin of natural language, and the relationship between learning and evolution. He is the director of the Center for the Conceptual Foundations of Science at the Parmenides Foundation, Pöcking, Germany. He is a professor of biology at Eötvös University in Budapest and a research professor at the Institute of Evolution of the Centre for Ecological Research.

Professor Szathmáry has recently published a seminal paper in PNAS on the prospects and threats of evolvable AI:

https://doi.org/10.1073/pnas.2527700123.
E-mail: szathmary.eors@gmail.com.

Balázs Szegedy is a Hungarian mathematician whose research concerns combinatorics and graph theory. Szegedy earned a master's degree in 1998 and a PhD in 2003 from Eötvös Loránd University in Budapest. His dissertation, supervised by Péter Pál Pálfy, was about group theory and was entitled "On the Sylow and Borel subgroups of classical groups". After temporary positions at the Alfréd Rényi Institute of Mathematics, Microsoft Research, and the Institute for Advanced Study, he joined the faculty of the University of Toronto Scarborough in 2006. He returned to the Rényi Institute in 2013. Balázs Szegedy is Member of the Hungarian Academy of Sciences.

Xu Wen is Professor of linguistics, and the former Dean of College of International Studies at Southwest University, Chongqing, China, where he lectures on semantics, cognitive linguistics, and sociolinguistics. His major research interests include cognitive linguistics, semantics, pragmatics, discourse analysis, syntax-semantics interface, cultural cognitive linguistics, second and foreign language education, and cognitive translation studies. His publications include *The Routledge Handbook of Cognitive Linguistics*, *COVID-19: Metaphor and Metonymy across Languages and Cultures*, and a great number of volumes and papers. He is the editor of the journal *Cognitive Linguistic Studies*, and *Asian-Pacific Journal of Second and Foreign Language Education*, and the book series co-editor of *Bloomsbury Studies in Language and Cognition*. He has organized the International Forum on Cognitive Linguistics since 2011. Xu Wen is President of China Cognitive Translation Society, and the vice president of China Cognitive Linguistics Association, and China Pragmatics Association. E-mail: xuwen@swu.edu.cn.

On Oct. 6, 2026, the Hungarian Academy of Sciences organizes an international conference on the topic of Artificial Intelligence from a philosophical point of view. Most of the speakers are not philosophers but represent various disciplines like psychology, medical science, communication theory, etc. So “philosophical” here means a return to the real task of philosophy: synthesizing for the educated public the ever new results of the different sciences. – The working language of the conference is English. The present bilingual booklet is a preliminary collection of the abstracts of the talks to be held. The booklet is edited by Kristóf Nyíri.

***

2026. okt. 6-ikán a Magyar Tudományos Akadémia konferenciát rendez, témája a mesterséges intelligencia filozófiai nézőpontból. Az előadók zöme nem filozófus, hanem változatos diszciplínákat képviselnek, úgymint pszichológiát, orvostudományt, kommunikációelméletet, stb. Azaz a „filozófia” itt visszatérést jelent a filozófia valódi feladatához: ahhoz, hogy a különböző tudományok új meg új fölfedezéseit a művelt nagyközönség számára szintetizálja. – A konferencia munkanyelve angol. Jelen kétnyelvű füzet a megtartandó előadások rövid kivonatainak előzetes gyűjteménye. A füzetet Nyíri Kristóf szerkesztette.

Kristóf [J. C.] Nyíri (1944), PhD, Dr. h.c., is full Member of the Hungarian Academy of Sciences, and a retired philosophy professor.

***

Nyíri Kristóf (sz. 1944), Dr. h.c., a Magyar Tudományos Akadémia rendes tagja. Nyugalmazott filozófiaprofesszor.