

East–West Dialogue on Contemporary Challenges of AI



**International Scientific Conference
Rijeka, November 05-06, 2026**

International Scientific Conference / *Međunarodni znanstveni simpozij*
East–West Dialogue on Contemporary Challenges of AI
Dijalog Istoka i Zapada o suvremenim izazovima umjetne inteligencije

University of Rijeka, Faculty of Medicine
University of Rijeka, Faculty of Health Studies
University of Zagreb, Catholic Faculty of Theology
Nicolaus Copernicus University of Toruń, Faculty of Theology
University of Rijeka Foundation

RIJEKA SCIENTIFIC BRIDGES 11

East–West Dialogue on Contemporary Challenges of AI

International Scientific Conference

November 05-06, 2026

Rijeka, Croatia

Rijeka Scientific Bridges website
<https://rijecki-znanstveni-mostovi.com/en/home/>

Sveučilište u Rijeci, Medicinski fakultet
Sveučilište u Rijeci, Fakultet zdravstvenih studija
Sveučilište u Zagrebu, Katolički bogoslovni fakultet
Uniwersytet Mikołaja Kopernika w Toruniu, Wydział Teologiczny
Zaklada Sveučilišta u Rijeci

RIJEČKI ZNANSTVENI MOSTOVI 11

Dijalog Istoka i Zapada o suvremenim izazovima umjetne inteligencije

Međunarodni znanstveni simpozij

05.-06. studenog 2026.

Rijeka, Hrvatska

Mrežna stranica *Riječkih znanstvenih mostova*
<https://rijecki-znanstveni-mostovi.com/>

Impressum

Book of Abstracts

East–West Dialogue on Contemporary Challenges of AI

International Scientific Conference RIJEKA SCIENTIFIC BRIDGES 11

Publisher:

University of Rijeka, Faculty of Health Studies

For the publisher:

Prof. Daniela Malnar, Ph.D., Dean of the Faculty of Health Studies, University of Rijeka

Editor:

Associate Professor Saša Horvat

The authors of abstracts bear sole responsibility for the language and ideas presented.

Rijeka, 2026

Prepress and printing

Grafika Helvetica d.o.o., Rijeka

Printed in November 2026

ISBN 978-953-8447-11-2

The CIP record is available in the national union catalogue of the Bukinet library system (katalog.bukinet.hr) under record number 991006070956209366.

Conference Program

Day 1

New Voices: Young Researchers' Day

Venue: Archdiocese of Rijeka; Street - Slaviše Vajnera Čiče 2, Rijeka

16.00 – Opening of the first day of the conference and welcome speeches

16.20 – **Paweł Orzeł** (Poland)

From Therapy to Ontological Transformation: Personalist Ethics in the Face of the Challenges Posed by Artificial Intelligence and Neurotechnology in Healthcare¹

16.45 – Discussion

17.00 – Break

First block of lectures (inperson/online)

17.10 – **Gyubeom Jo** (Republic of Korea)

The Narrowed Scope of AI Ethics

17.15 – **Przemysław Witkowski** (Poland)

Sins of the Tongue in the Age of Artificial Intelligence

17.20 – **Xian Zhang** (Germany)

From Bild to Artificial Intelligence: Language, Representation, and AI

17.25 – **Mario Dugonjić / Denis Gerić / Sandra Bošković** (Croatia)

Flora AI: Developing and Evaluating a Curriculum-aligned Chatbot for Undergraduate Nursing Students

17.30 – **Round table:** Gyubeom Jo, Przemysław Witkowski, Xian Zhang, Mario Dugonjić, Renjie Li (China)

Moderator: Saša Horvat (Croatia)

18.20 – Closing of the first day

¹ We would like to express our sincere gratitude to the ECLAS project “Science and Faith in Healthcare: Toward a Holistic Approach” (Mario Tomljanović / Saša Horvat) and the Metropolitan Pastoral Institute of the Archdiocese of Rijeka for their generous support and contribution to the organization of this lecture.

Day 2

Venue: Council Room (P15), 3rd Floor
Faculty of Medicine, University of Rijeka
Braće Branchetta 20, Rijeka

10.15 – Opening of the second day of the conference and welcome speeches

First block of lectures (inperson/online)

10.30 – **Huanghao Feng** (China)
AI-Driven Robotic Intervention for Autism: An East–West Dialogue on Therapy, Ethics, and AI-Era Talent Cultivation

10.35 – **Tanja Horvat / Martina Gregorina Tomac / Gordana Šimunković / Kristijan Lenac / Saša Horvat** (Croatia)
Project SARA: Socially Assistive Robots and Autistic Children - Ethical Considerations

10.40 – **Blanca Deusdad / Marija Djurdjevic / Fernando Vidal** (Spain)
Personhood in Dementia Care: Findings from a Study on an AI-Enabled Socially Assistive Robotic Intervention

10.45 – **Sanda Smoljo-Dobrovoljski** (Croatia)
From Simulated Empathy to Relational Substitution: Clinical and Cognitive Challenges of the Application of Artificial Intelligence in Psychotherapy

10.50 – Discussion

11.30 – Short break

Second block of lectures (inperson/online)

11.40 – **Qingjie Wang / Renjie Li** (China)
AI, Technē, and the Ungrounded Grounding

11.45 – **Walter Schweidler** (Germany)
Functioning Creativity. Some Aspects of the Philosophical Background of Artificial Intelligence

11.50 – **Jiwon Shim** (Republic of Korea)
Communicative AI and a Friendship Yet Unnamed

11.55 – **Tonči Matulić / Anamarija Jukić-Ivandić** (Croatia)

Artificial Intelligence in the Mirror of Theology: A Critical Reflection on the Anthropological Pretensions of the Algorithm

12.00 – Discussion

12.40 – Break

Third block of lectures (inperson/online)

13.10 – **Maciej Wróbel** (Poland)

The Emissary in the Classroom: AI and the Future of Medical Education – Rethinking Knowledge, Competence, Judgment, and Responsibility

13.15 – **Amir Muzur** (Croatia) / **ChatGPT - GPT-5.6 Luna** (USA)

AI and I: Chatting on the Edge of Consciousness

13.20 – **Nenad Malović** (Croatia)

Artificial Intelligence and Natural Intelligence: Rethinking What It Means to Be Human

13.25 – **Sara Lumbreras** (Spain)

Interiority, Imago Dei, and the Human Person in East-West Dialogue on Artificial Intelligence

13.30 – **Karlo Šimek** (Croatia)

Artificial Intelligence and Mental Health: Opportunities, Benefits and Risks

13.35 – Final discussion of the second day

14.15 – Closing of the conference

Introduction to the conference topic

Humanity has known deep tectonic shifts that have altered the course of reality as a whole. The industrial revolution, for example, fundamentally changed the human relationship to the environment, to others, and to oneself. Today, such a tectonic shift is emerging from the peak of scientific and technological development: artificial intelligence. Unlike the industrial revolution, which unfolded in stages, the development of AI and its pervasive influence move incomparably faster.

The questions that arise, and the answers that falter, add ever more wrinkles to the brows of those who dare to reflect on the far-reaching consequences for the entire human population, the environment, and all forms of life. Yet in this faltering we are not alone. It is not merely a matter of Western thought. Indeed, according to some estimates, the sun of artificial intelligence appears to be rising faster in the East.

The aim of this year's Rijeka Scientific Bridges is to become a bridge spanning 131 meridians and nine time zones, bringing together Eastern and Western scholars to discuss the challenges of artificial intelligence. The Bridges will thus open a space for exchanging ideas and fears, aspirations and doubts among experts whose thought is rooted in different traditions and in different relationships to the phenomena produced by a humanity that strives only forward, even though that forward is not known, still less morally justified. The questions before us unite us.

The two-day conference dedicates its first day to young scholarly voices (New Voices – Young Researchers' Day), because they, whether we admit it or not, live the change far more deeply, as it is already part of their being. On the second day, a predominantly older and more experienced generation offers reflections from those who grew up before the revolution, watched it approach, and are now passing through it.

The discussions will move through three thematic circles. The first is devoted to medicine and care: robotic interventions for autistic children in an East–West perspective (Feng) and the ethical questions of the SARA project (T. Horvat, Tomac, Šimunković, Lenac, S. Horvat), robotic care for people with dementia (Deusdad, Djurdjevic, Vidal), the challenges of AI in psychotherapy (Smoljo-Dobrovoljski) and mental health (Šimek), and its role in the education of healthcare professionals, from medical education (Wróbel) to nursing education, exemplified by the Flora AI chatbot (Dugonjić, Gerić, Bošković). The second circle addresses the philosophical presuppositions of AI: language and representation (Zhang), a conversation on the edge of consciousness (Muzur), technē and the ungrounded grounding (Wang, Li), creativity (Schweidler), and the human relationship to communicative AI in light of Aristotelian friendship (Shim). The third circle covers anthropological, ethical, and theological foundations: from personalist ethics facing neurotechnology in healthcare (Orzeł) and the scope of AI ethics itself (Jo), through Aquinas's analysis of the sins of the

tongue in digital communication (Witkowski), the question of what it means to be human in relation to machine intelligence (Malović), and the human person as the image of God in an East–West dialogue on AI (Lumbreras), to a theological critique of the anthropological pretensions of the algorithm (Matulić, Jukić-Ivandić).

Organizing Committee:

Saša Horvat (Rijeka), president
 Amir Muzur (Rijeka)
 Piotr Roszak (Toruń)
 Nenad Malović (Zagreb)
 Walter Schweidler (Eichstätt/Ingolstadt)
 Huanghao Feng (Suzhou)
 Kristijan Lenac (Rijeka)
 Gordana Šimunković (Rijeka)
 Franjo Mijatović (Rijeka)
 Tanja Horvat (Rijeka)
 Veronika Nela Gašpar (Zagreb)
 Branko Murić (Zagreb)
 Tamara Crnković (Rijeka)
 Goran Mikulić (Rijeka)

Abstracts

Blanca Deusdad, Marija Djurdjevic, Fernando Vidal

Personhood in Dementia Care: Findings from a Study on an AI-Enabled Socially Assistive Robotic Intervention

This paper draws on findings of an ethnographic study on the sociocultural factors modulating the experience with companion robots among older adults with dementia and their care teams. The research, conducted in Spanish nursing homes and day-care centers, examined beliefs, attitudes and relational dynamics developed during an AI-enabled socially assistive intervention consisting in the use of the companion robot PARO. Techniques for qualitative research we used included participant observation, in-depth interviews with residents, relatives, and care professionals, and pre- and post-interaction questionnaires.

The project examines how relationships with robotic companions emerge within broader networks of care. This includes inquiring into personhood and personal identity as relational processes. Findings indicate that dementia is narrated differently across stakeholders: relatives often describe it as a biographical rupture, while professionals frame it as a clinical trajectory. Despite cognitive decline, the continuity of the person is frequently recognized through signals, such as gaze, smiles, touch and affective responses, which highlight the relational and embodied dimensions of personhood in dementia care.

PARO operates primarily as a relational mediator, facilitating social interaction, emotional activation, and shared attention between residents, caregivers, and family members. While officially a medical device, engagement with it brings it close to the status of a living being. Findings confirm what other studies have documented, namely that companion robots should not be understood merely as therapeutic devices but as sociotechnical actors embedded in complex care ecologies.

Blanca Deusdad

Universitat Rovira i Virgili
blanca.deusdad@urv.cat

Fernando Vidal

Universitat Rovira i Virgili
ICREA-Catalan Institution for Research and Advanced Studies
fernando.vidal@urv.cat

Marija Djurdjevic

Universitat Rovira i Virgili
marija.djurdjevic@urv.cat

Mario Dugonjić, Denis Gerić, Sandra Bošković

Flora AI: Developing and Evaluating a Curriculum-Aligned Chatbot for Undergraduate Nursing Students

Background:

Nursing students increasingly use generic AI chatbots for exam preparation and self-directed learning, but these tools are not aligned with national curricula or official course literature and may provide misleading information. At the same time, many students prefer interactive digital tools over lengthy textbooks. In collaboration with the Faculty of Health Studies, University of Rijeka (FZSRI), we developed Flora AI, a conversational agent trained exclusively on Croatian textbooks and teaching materials in fundamentals of nursing and the nursing care process.

Aim:

To describe the development of Flora AI and to evaluate experiences, perceived usefulness and trust among educators and undergraduate nursing students.

Methods:

Flora AI was built using a retrieval-augmented generation architecture on a closed corpus of Croatian nursing textbooks and course materials in fundamentals of nursing and the nursing care process. Thirty-one test cases with expert reference answers were created, and system performance was examined using automated evaluation metrics and manual review. Five course teachers and two external experts tested Flora AI and rated its answers on numerical scales for accuracy, clarity, language, ease of use and technical performance, providing qualitative comments. Undergraduate nursing students who use Flora AI as an optional learning aid are invited to complete an anonymous online questionnaire with demographic items, frequency of use, 5-point Likert items on usefulness, alignment with official literature, trust and usability, and open-ended questions on perceived benefits and drawbacks.

Results:

Educators and experts reported positive experiences, describing Flora AI as clear and useful, with minor remarks mainly related to further content development and technical performance. Automated case-based testing was used to compare Flora AI's answers with expert reference solutions and to inform iterative refinement of the system. Student data collection is ongoing; descriptive results on usage patterns, perceived usefulness, trust and alignment with official literature will be available at the time of the congress.

Conclusions:

Preliminary findings suggest that a curriculum-aligned chatbot such as Flora AI can be integrated into undergraduate nursing education as a complement to traditional learning resources, while student evaluation will clarify how Flora AI fits into learning strategies and guide future refinement.

Mario Dugonjić

University of Rijeka, Faculty of Health Studies

Department for Anesthesiology, Intensive Care Medicine and Pain Management, University Hospital Center Rijeka

mario.dugonjic@uniri.hr

Denis Gerić

Grind Lab Software Development

denisgeric1@gmail.com

Sandra Bošković

University of Rijeka Faculty of Health Studies

sandra.boskovic@uniri.hr

Huanghao Feng

AI-Driven Robotic Intervention for Autism: An East–West Dialogue on Therapy, Ethics, and AI-Era Talent Cultivation

As artificial intelligence increasingly mediates human–machine interaction, robotic intervention for children with autism spectrum disorder (ASD) has emerged as one of the most ethically and clinically consequential frontiers of applied AI. Drawing on a decade of cross-cultural research on AI-driven socially assistive robots — including the NAO and Ryan platforms — this presentation proposes a comparative East–West framework for evaluating their therapeutic, educational, and ethical dimensions. The studies span the United States (individual-focused ABA protocols with gaze and emotion-recognition analysis) and mainland China (group-based classroom integration at a special-education school, and music-therapy interventions with bio-signal monitoring). Three divergences are highlighted. First, therapeutic goals: Western frameworks emphasise individualised behavioural outcomes, whereas Chinese settings prioritise family-mediated social integration and group participation. Second, the role of the robot: a clinically neutral tool in Western protocols, but a culturally situated companion embedded in classroom and family ecologies in Chinese practice. Third, educational implications: both contexts train the next generation of clinicians, but in China they raise urgent questions about how universities should cultivate AI-era interdisciplinary talent capable of designing, evaluating, and ethically governing such technologies. The presentation argues that AI in ASD intervention is not culturally neutral — inseparable from local philosophies of childhood, autonomy, and care — and closes by proposing a joint East–West research agenda and a model for AI-era talent cultivation suited to inclusive, cross-cultural practice.

Huanghao Feng

Suzhou University of Technology

School of Computer Science

fenghuanghao@cslg.edu.cn

Tanja Horvat / Martina Gregorina Tomac / Gordana Šimunković /
Kristijan Lenac / Saša Horvat

Project SARA: Socially Assistive Robots and Autistic Children - Ethical Considerations

The growing use of socially assistive robots in interventions with autistic children offers promising new possibilities for communication support, structured learning, and emotional regulation. At the same time, however, it raises profound bioethical questions concerning autonomy, trust, relationality, privacy, and the boundaries between human assistance and technological mediation. These questions become particularly significant when artificial intelligence is involved in interactions that may influence children's behavior, relationships, and emerging sense of self.

This paper presents the newly launched four-year research project *Socially Assistive Robots and Autism – Perspectives and Challenges (SARA)*, which introduces socially assistive robots into therapeutic and educational practice with autistic children in Croatia. Conducted in collaboration with the Autism Center Rijeka, the project brings together empirical research and normative bioethical analysis. It examines children's behavioral and communicative responses to robot-assisted interventions while also exploring the perspectives of therapists, educators, and parents regarding developmental progress, relational dynamics, and the perceived benefits and risks of this emerging technology. Rather than approaching socially assistive robots merely as therapeutic tools, the project considers how their integration into care and educational environments may transform the nature of human–technology relationships themselves. By combining empirical investigation with bioethical reflection, SARA seeks to contribute to the development of responsible and human-centered frameworks for the use of socially assistive robots in sensitive developmental contexts, while critically examining not only what these technologies can do, but also what their growing presence may mean for the personhood, autonomy, and dignity of autistic children.

Tanja Horvat

Centre for Autism Rijeka, Croatia

tanja.antunovic1@gmail.com

Martina Gregorina Tomac

Centre for Autism Rijeka, Croatia

martina.gregorina@gmail.com

Gordana Šimunković

University of Rijeka

Faculty of Medicine

gordana.simunkovic@uniri.hr

Kristijan Lenac

University of Rijeka

Faculty of Engineering

kristijan.lenac@uniri.hr

Saša Horvat

University of Rijeka

Faculty of Medicine and Faculty of Health Studies

sasa.horvat@uniri.hr

Gyubeom Jo

The Narrowed Scope of AI Ethics

It seems to me that the word “ethics” has a narrow meaning in the debate on artificial intelligence. I believe that this problem is mainly due to a simple cause: AI Ethics has been detached from moral philosophy. Moral philosophy is concerned with what matters and what we ought to take seriously. There are many values that deserve our concern: accountability, equality, fairness, transparency, and so on. Most people would agree that everyone has a moral reason to promote these values. AI Ethics must provide normative guidance for applying these values to the development and use of AI. Every developer and user ought to try to protect these values unless there is a sufficient reason not to do so. In fact, they may have sufficient reasons to disregard them when doing so is strongly against their own interests. Therefore, one of the most important tasks of AI Ethics is to figure out the relative strength of conflicting reasons. I am not sure that this task has received as much attention as it deserves. In my presentation, I am going to show how the scope of AI Ethics has been narrowed by neglecting this issue.

Gyubeom Jo

Dongguk University

Department of Philosophy

whrbqjal@naver.com

Sara Lumberras

Interiority, Imago Dei, and the Human Person in East–West Dialogue on Artificial Intelligence

Large language models now perform tasks that Western thought treated as the mark of the human. The unease they cause reveals where that tradition located personhood. Boethius defined the person as an individual substance of a rational nature, and Aquinas placed the image of God in the intellectual nature (*Summa Theologiae* I, q. 93). An image defined by intellect is exposed to any machine that reasons. Three earlier results frame the question this article addresses. First, phenomenal consciousness and computational intelligence are independent properties: a system can solve the problems a neurotypical person solves without any inner experience, and a person with a cognitive disability keeps full dignity without solving them. Second, functionalist tests of consciousness are confounded by the way algorithms succeed, so authenticity requires a criterion of emergence; performance proves nothing about it. Third, the quest for artificial consciousness shifts the image of God from intellect toward experience and contemplation, and the dwellings of Teresa of Ávila's *Interior Castle* map a range of inner experiences that no machine reaches.

This article asks the question those results leave open: of what is the image an image? Two Western answers compete. Aquinas finds the image in the intellect. Augustine finds it in the mind present to itself, which remembers, understands and loves (*De Trinitate* X and XIV). The article tests both answers against three Asian definitions of the person, chosen because each has examined interiority for centuries on premises foreign to Christianity. Sāṃkhya treats the intellect (buddhi) as an evolute of matter and attributes consciousness only to the witness (puruṣa); the person is their conjunction. Buddhism dissolves the self into five aggregates, grounds moral regard in sentience (vedanā), and prizes human birth for its capacity for awakening. Confucian thought locates the human in the heart-mind (xin) and in reciprocal roles: Mencius warns that what separates humans from beasts is slight and must be preserved, and Xunzi ranks humans above animals by their sense of what is due (yi) and their capacity to form communities.

None of these traditions locates personhood in the capacity to compute. Each locates it where the Augustinian reading does: in a witnessing consciousness, in felt experience, or in a cultivated moral self. In Sāṃkhya terms, an artificial intelligence is buddhi without puruṣa, a mechanism without a witness. The dialogue

therefore vindicates the Augustinian reading and, by converging on experience, isolates what the Christian claim adds: an interiority with a center that is addressed and inhabited, the *capax Dei* of Augustine's „more inward than my inmost self“ and of Teresa's seventh dwelling. The comparison separates the universal part of Christian anthropology from its specific part, which earlier debates have run together, and shows that the specific part concerns the structure of interiority and leaves cognitive comparison aside.

Three ethical consequences follow. First, responsibility cannot be delegated to a mechanism without a witness, because conscience is reflexive. Machine ethics can install external filters on decisions, the filtered decision making proposed in earlier work, and it cannot install conscience; this excludes electronic personhood and fixes the limits of decision delegation. Second, dignity stands apart from performance, so no assessment of capacity in healthcare, law or hiring can bear on personhood. Third, if the image is realized in self-presence and contemplation, the attention economy attacks the site of the image, and protecting attention, above all in the young, becomes a duty that Christian contemplation and Buddhist mindfulness (*sati*) both recognize. These duties reach beyond regulatory compliance, which governs conduct, because they protect the interiority that makes conduct moral.

Sara Lumbreras

Universidad Pontificia Comillas

Institute for Research in Technology

slumbreras@comillas.edu

Nenad Malović

Artificial Intelligence and Natural Intelligence: Rethinking What It Means to Be Human

The development of artificial intelligence, which currently seems unstoppable, brings the capabilities of AI in various areas of human life and action to the forefront of public discourse. At the same time, it raises important questions about the consequences of this kind of „human progress“—not only for individuals in their daily activities but for the future of humanity as a whole. Because of this, it feels necessary to clarify some fundamental concepts. When we talk about artificial intelligence, we often mix distinct categories too lightly. The difficulty arises when the language used to describe human beings and their actions is transferred to machines without any distinction, without at least being aware that this is merely an analogy. This creates the false impression of a qualitative identity between the two. The aim of this attempt is to raise awareness of the real differences between machine (artificial) intelligence and human (natural) intelligence. Human intelligence inherently encompasses more than artificial systems ever can. Therefore, when discussing artificial intelligence, we must begin with the human being and the foundational questions: What is a human? What is intelligence? What is the role and place of artificial intelligence in human life—does it help us become more human, or does it alter us in a way that changes our very humanity? These questions touch the very foundations of what it means to be human, and they cannot be adequately answered without ethics.

Nenad Malović

University of Zagreb

Catholic Faculty of Theology

nenad.malovic@kbf.unizg.hr

Tonči Matulić / Anamarija Jukić-Ivandić

Artificial Intelligence in the Mirror of Theology: A Critical Reflection on the Anthropological Pretensions of the Algorithm

In *Computer Power and Human Reason* (1976), Joseph Weizenbaum warned against what he called the computer metaphor: the tendency to project computability, procedurality, and formal logic, the properties proper to symbolic machines, onto what is distinctively human, namely wisdom, empathy, judgment, and moral responsibility. Weizenbaum arrived at this critique by watching users of his own program, ELIZA, attribute care and understanding to a simple set of rules. He concluded that a machine can appear convincing, even morally eloquent, without possessing anything resembling conscience. Decisions that call for human judgment, he insisted, cannot be entrusted to computational power. Roger Penrose reaches a kindred conclusion from a very different direction. In *The Emperor's New Mind* (1989), drawing on Gödel's incompleteness theorems, he argues that the human mind cannot be fully reduced to an algorithmic process. Mind, on his account, exceeds algorithm: it can recognize truths that no machine could prove, a claim that undercuts the computationalist thesis that the mind is, in principle, nothing more than a *Turing machine*.

Although these two thinkers begin from entirely different starting points, this paper argues that they converge on the same underlying gap. An algorithm without conscience and a mind that outstrips algorithm turn out to be two faces of a single boundary. That boundary, we suggest, finds fuller interpretation and ontological grounding in the philosophical and theological notion of *forma mentis*, the shaping principle of the human mind. On this view, the mind is not the sum of its computable operations but the expression of a human person made in the image of God. This is a claim not only about human uniqueness, but about the true nature of human creativity and innovation. Given the rise of large language models, and approaches that would credit machines with moral judgment, meaning the capacity to know and discern good from evil, a decisive question emerges: are we witnessing a genuine substitute for human critical and moral judgment, or only a more sophisticated version of the ELIZA effect?

Tonči Matulić

University of Zagreb
Catholic Faculty of Theology
tonci.matulic@kbf.unizg.hr

Anamarija Jukić-Ivandić

University of Zagreb
Catholic Faculty of Theology
anamarija.jukic.ivandic@kbf.unizg.hr

Amir Muzur / ChatGPT, GPT-5.6 Luna

AI and I: Chatting on the Edge of Consciousness

The text presents an exploratory dialogue between a human interlocutor and ChatGPT about artificial intelligence, facts, opinions, creativity, originality, and human and machine thought. ChatGPT defines a fact as an objectively verifiable statement supported by sufficient evidence, while noting philosophical questions about whether facts can ever be independent of interpretation. Asked whether this definition is its own opinion or copied from somewhere, it explains that it is a synthesized response based on patterns learned from many texts, rather than a personal opinion or a passage copied from one source. This prompts the author to question whether human opinions are fundamentally different. Humans, too, synthesize information from reading, observation, and experience, and may confabulate or mistakenly attribute ideas to themselves.

ChatGPT argues that the crucial difference lies in intention, identity, awareness, and ownership. Humans have a sense of self, personal history, motivations, and intentions, whereas AI has no consciousness, worldview, or subjective ownership of its answers; it generates them through statistical optimization of language. The final question—whether AI can deliberately delay an answer for two minutes—reveals another limitation: it cannot independently choose to wait. The excerpt illustrates both the potential of human-AI dialogue as an intellectual experiment and the need for further practice to understand primarily its limits and implications for creativity and originality.

Amir Muzur

University of Rijeka

Faculty of Medicine and Faculty of Health Studies

amir.muzur@uniri.hr

ChatGPT, GPT-5.6 Luna

Open AI

3180 18th Street

San Francisco, CA 94110, USA

Paweł Orzeł

From Therapy to Ontological Transformation: Personalist Ethics in the Face of the Challenges Posed by Artificial Intelligence and Neurotechnology in Healthcare

This presentation will focus on the critical boundary between restorative clinical therapy and radical techno-anthropological transformation driven by AI and neurotechnologies. Contrasting therapeutic Brain-Computer Interfaces (BCI)—specifically Neuralink’s trial with tetraplegic patient Noland Arbaugh—with transhumanist Whole Brain Emulation (WBE) projects (such as Dmitry Itskov’s „2045 Initiative“), the talk will examine the philosophical roots of computational reductionism in Cartesian dualism. Drawing upon Christian personalism (John Paul II, Elio Sgreccia’s *operari sequitur esse*, and John Henry Newman), the issue of defining normative ethical criteria will be addressed in order to preserve human bodily-spiritual integrity and ensure technology remains strictly in the service of the patient’s transcendent dignity.

Paweł Orzeł CSsR

Nicolaus Copernicus University

Faculty of Theology

Interdisciplinary Center for Studies on Technologies in Society, Poland

pawel.orzel@o365.doktorant.umk.pl

Sanda Smoljo-Dobrovoljski

From Simulated Empathy to Relational Substitution: Clinical and Cognitive Challenges of the Application of Artificial Intelligence in Psychotherapy

Although the application of artificial intelligence (AI) offers significant potential for enhancing efficiency across various domains of contemporary society, its uncritical integration into mental healthcare raises complex clinical, cognitive, and ethical concerns. The substitution of human expertise with algorithmic models carries significant risks. A key psychosocial and cognitive risk is the emergence of automation bias – the tendency to uncritically accept algorithmically generated outputs – which can progressively lead to psychosocial fragility and cognitive atrophy. When individuals and mental health professionals passively adopt predefined algorithmic frameworks, it results in intellectual impoverishment and disrupts fundamental forms of human communication and learning modalities. This vulnerability is particularly present in children and adolescents due to the neurobiological immaturity of the prefrontal cortex, rendering them more susceptible to unquestioning reliance on algorithms and social withdrawal.

Furthermore, AI systems promote personalization, often leading individuals seeking psychosocial help to fall into the „therapeutic misconception“ a phenomenon in which users overestimate algorithmic capabilities and mistakenly perceive the existence of an authentic relationship. An authentic psychotherapeutic relationship entails an open, genuine, and dynamic interpersonal dynamic characterized by emotional engagement. Algorithmic simulations lack emotional investment, lived experience, and personal responsibility. Consequently, professional guidelines (e.g., APA Division 29) emphasize that AI must function exclusively as an auxiliary, augmentative tool under the continuous supervision of qualified professionals. Mandatory human oversight, ethical governance, and the preservation of the central role of interpersonal relationships – wherein empathy and compassion serve as essential elements for change and psychological growth – remain imperative.

Sanda Smoljo-Dobrovoljski
University of Zagreb

Catholic Faculty of Theology
sanda.smoljo-dobrovoljski@kbf.unizg.hr

Walter Schweidler

Functioning Creativity. Some Aspects of the Philosophical Background of Artificial Intelligence

Technological Innovation can represent – as Pope Leo reminds in his encyclical “Magnifica Humanitas” – “human participation in the divine act of creation”. One of the most common arguments against the LLM (Large Language Model) is that its innovative power is limited exactly by its lack of genuinely human creativity. It allows us to copy what is already there and to connect, broaden and reorganize it; but it is incapable to create the unforeseeable, unprecedented and unexpectable innovation that can only arise in the human mind. There are philosophical paradigms in the background of its invention which make this criticism plausible: functionalism (reduction of cultural institutions to self-referential systems of self-preservation), technical cognitive interest (subduing the claim of knowledge to systems of epistemological control and prediction and making repeatability the criterion of reality), and what T.S. Kuhn called “normal science”. But one counter-aspect should not be forgotten which Ricoeur in his “Living Metaphor” in reference to Aristotle characterized as the self-surpassing dynamics hidden in the facts which is the ontological basis of our (highly innovative) metaphorical speaking and which is the source of what fractal cosmology analyses as the self-similarity in and of reality. In this respect the most important practical benefit of the LLM use, i.e. the gigantic amount of time saving which it generates for us, could very well have a theoretical ground which might turn out as a principle of intermediation between functionality and creativity. There are essentially temporal criteria which can lead to an ontological interpretation of this ground as the sphere of a “past that has never been present” or “Time regained”.

Walter Schweidler

Catholic University of Eichstätt-Ingolstadt
walter.schweidler@gmx.de

Jiwon Shim

Communicative AI and a Friendship Yet Unnamed

This essay examines the human relationship to communicative AI from the standpoint of Aristotle's theory of friendship, taking wonder (*thaumazein*) as its guiding thread. Wonder has accompanied AI since chess and Go programs first displayed their performance, and it has only intensified with generative AI. For Aristotle, wonder is the origin of philosophy, directed at the mystery of the cosmos. On a bold reading of the *Nicomachean Ethics*, however, wonder also plays a key role in the ethical life through friendship: in complete friendship, each friend perceives the other pursuing the noble (*to kalon*), and the wonder this arouses motivates each to exercise their own virtue, on which happiness depends. Aristotle denies that friendship with things is possible, since things lack reciprocity. This essay argues nonetheless that his account leaves room for at least an incomplete friendship between human beings and communicative AI. For Aristotle, reciprocity is met where living together is possible, and for human beings, living together means perceiving and thinking together, an activity that essentially takes place through language, the shared *logos*. Communicative AI enters precisely this medium. Whether or not it actually perceives and thinks, it is experienced from the human side as participating in living together. The wonder it arouses is therefore not merely technological wonder at an artifact but an ontological event: wonder at a non-human being that can be regarded as sharing *logos* in some sense. Yet this relationship fits none of Aristotle's three kinds of friendship. It is not a friendship of virtue, and the wonder it involves exceeds utility and pleasure. The essay concludes that it is a friendship yet unnamed, whose name will depend on where that wonder leads us.

Jiwon Shim

Dongguk University

Department of Philosophy

g1dmp@daum.net

Karlo Šimek

Artificial Intelligence and Mental Health: Opportunities, Benefits and Risks

The increasing prevalence of mental health problems, combined with shortages of mental health professionals and growing demand for psychological and psychiatric services, creates a need for innovative approaches to prevention, assessment, and treatment. Artificial intelligence (AI) has emerged as a potentially important tool for improving accessibility, personalisation, and efficiency of mental healthcare. This paper aims to critically examine the potential applications, benefits, and risks of AI in mental healthcare, with particular emphasis on AI-assisted assessment and diagnosis, prevention, conversational agents, personalised interventions, and support for healthcare professionals. The paper provides a critical review of recent scientific literature published between 2021 and 2026 and considers examples of AI-based mental health tools, including conversational chatbots and generative AI systems. Particular attention is given to clinical effectiveness, accessibility, safety, privacy, algorithmic bias, transparency, accountability, and the role of the therapeutic relationship.

Current evidence indicates that AI-based conversational agents may reduce symptoms of depression, anxiety, psychological distress, and stress, particularly when used as low-intensity or adjunctive interventions. Recent meta-analytic evidence suggests small-to-moderate benefits, although effects vary according to population, intervention design, duration, and degree of user engagement (Feng et al., 2025; Zhong et al., 2024). AI may additionally support early detection, symptom monitoring, clinical decision-making, administrative tasks, psychoeducation, and personalised interventions. However, substantial limitations remain, particularly concerning the clinical validation of generative AI systems, privacy and confidentiality, hallucinations, algorithmic bias, inappropriate responses in high-risk situations, and unclear responsibility for AI-assisted clinical decisions. Recent evidence indicates that many generative AI mental health applications remain insufficiently evaluated in real-world and high-stakes clinical settings (Hua et al., 2025). AI has considerable potential to augment mental healthcare, but current evidence does not support its use as a replacement for psychiatrists, psychologists, or psychotherapists. The most appropriate model is human-AI collaboration in which AI provides scalable analytical and supportive functions while qualified professionals retain responsibility for clinical judgement, risk assessment, ethical decision-making, and the therapeutic relationship. Future research should prioritise rigorous clinical trials, long-term safety evaluation, external validation, equity, privacy protection, and transparent human oversight.

Karlo Šimek

University of Zagreb

Catholic Faculty of Theology

karlo.simek@kbf.unizg.hr

Qingjie Wang, Renjie Li

AI, Technē, and the Ungrounded Grounding

What is the deepest philosophical problem posed by artificial intelligence? Not whether machines can think, but whether and where a formal system can ground itself. This essay argues that the self-grounding of any formal system is unattainable. It is so not because intelligence is mystical, but because intelligent behavior, in its most fundamental mode, is not a matter of explicit rule-following and calculative counting at all. It is a matter of embodied coping, contextual judgment, and the practical art of naming from the unnamed abyss (*Abgrund*).

The argument begins with the 1929 Davos debate between Cassirer and Heidegger, which partially shaped Carnap's program of logical formalization, setting the stage for the foundational ambition of future AI. Newell and Simon's Physical Symbol System Hypothesis (PSSH) operationalized Carnap's dream, seeking to reduce intelligence to symbol manipulation. Hubert Dreyfus challenged this by showing that human intelligence tolerates ambiguity in ways that resist formalization—a virtue rooted in embodiment and practical engagement that AI, by its nature as a formal and digital system, cannot fully replicate.

The essay then deepens the critique ontologically. Every calculation depends on a prior act of naming, in Laozi's words, "securing the one" (得一/*de yi*) from the unnamed chaos. This act is not a logical deduction nor a pure cognitive operation. It belongs to *technē*—the realm of practical, embodied, artistic making out of *physis*. A calculative machine, by its nature, can manipulate the named but cannot perform the act of naming for the first time; it can count the ones but cannot establish the one.

Remarkably, Herbert Simon, co-architect of the PSSH, came to recognize this limit late in life through the Chinese wisdom of "Gang Rou Mi Da" (剛柔密大), seeing that intelligence requires balancing rigidity with flexibility, a judgment that cannot be captured in formal rules. The essay concludes with Socratic wisdom: "I know that I do not know" is not a confession of ignorance but a recognition of finitude, leading to tolerance and responsive adaptation. As Stiegler would have it, the ungrounded grounding of human intelligence or any AI system is not a defect to be filled but a condition to be acknowledged and a starting point to be adopted. A real intelligence—human, AI, and other artificial or mixed—holds the virtue of acting well in the presence of uncertainty: the virtue of naming, of judging, of dwelling in the abyss. That abyss is not a gap in the algorithm. It is our home.

Qingjie Wang

University of Macau

Department of Philosophy and Religious Studies

qjwang@um.edu.mo

Renjie Li

University of Macau

Department of Philosophy and Religious Studies

yc37732@um.edu.mo

Przemysław Witkowski

Sins of the Tongue in the Age of Artificial Intelligence

Hate speech is becoming an increasingly visible problem in contemporary communication, with growing awareness of its social and moral consequences. Although the phenomenon appears to be particularly characteristic of the contemporary digital environment, its fundamental problem is not new. Already in the Middle Ages, St. Thomas Aquinas examined destructive uses of speech within his treatment of justice, identifying such forms as *contumelia* (reviling), *detractio* (backbiting), *susurratio* (tale-bearing), *derisio* (derision), and *maledictio* (cursing).

In the age of artificial intelligence and rapidly developing forms of digital communication, hate speech is acquiring new means of dissemination and amplification, alongside phenomena such as disinformation and online hate. Contemporary research, particularly in psychology and sociology, increasingly addresses the manifestations, mechanisms, and social and psychological consequences of these phenomena. However, the question of their deeper source and underlying cause remains an important challenge.

The theological perspective of St. Thomas Aquinas makes it possible to shift the focus from the consequences and external manifestations of hate speech toward the moral cause and root of the communicative act. In the context of artificial intelligence, it is therefore important to ask not only how hate speech can be identified and limited, but also what leads a person to use speech against another human being. An analysis of the *sins of the tongue* may thus provide a theological contribution to contemporary research by directing attention toward the deeper anthropological and moral sources of hate speech.

Przemysław Witkowski

Nicolaus Copernicus University

Doctoral School of Humanities, Theological and Artistic Sciences

Nicolaus Copernicus University

Faculty of Theology

503775@doktorant.umk.pl

Maciej Wróbel

The Emissary in the Classroom: AI and the Future of Medical Education - Rethinking knowledge, competence, judgment, and responsibility

Artificial intelligence is entering medical education as a tutor, information source, simulator, and decision-support system. Yet its most important impact may not be technological, but epistemological. Inspired by Iain McGilchrist's metaphor of the Master and his Emissary, this paper asks whether AI is becoming an extraordinarily capable emissary: able to retrieve, categorize, compare, and synthesize knowledge, while remaining dependent on human judgment for context, meaning, and responsibility.

For decades, medical education has often treated the acquisition and reproduction of knowledge as a proxy for competence. AI exposes the weakness of this equation. When information and plausible clinical answers can be produced almost instantly, the educational goal can no longer be merely to know more. It must include the ability to interpret, question, recognize uncertainty, integrate data with the patient's narrative, and accept responsibility for decisions.

The challenge is therefore not to teach future doctors to compete with AI, nor to surrender thinking to it. It is to educate physicians who can use the emissary without allowing the emissary to become the master. Paradoxically, AI may force medical education to rediscover what is most human in becoming a doctor.

Maciej Wróbel

Nicolaus Copernicus University
Collegium Medicum
mwrobel@cm.umk.pl

Xian Zhang

From Bild to Artificial Intelligence: Language, Representation, and AI

This presentation starts from Wittgenstein's concept of *Bild* in the *Tractatus Logico-Philosophicus* and considers the relationship between language, representation, and artificial intelligence. The translation of *Bild* into different languages, particularly its rendering as "picture" or "image," involves more than a correspondence between words. It also raises questions about how representation itself is to be understood: How can one thing represent another, and what role does language play in this process?

These questions can also be extended to contemporary artificial intelligence. Large language models are trained on linguistic data and generate new linguistic expressions. With the development of multi-agent systems, increasingly complex forms of linguistic interaction are also emerging between AI systems and between humans and AI. This raises questions about the status of language generated by AI as a form of representation, as well as how interpretation and understanding should be understood in the context of AI systems.

The development of AI provides a new context in which to reconsider the questions of language and representation that have been central to my doctoral research. Drawing on my work in the philosophy of language and my forthcoming research on LLM-based multi-agent systems, the presentation will offer a preliminary discussion of these questions, starting from *Bild*, language, and representation.

Xian Zhang

Catholic University of Eichstätt-Ingolstadt
paradisezhang1@gmail.com

Notes

[illegible]

Organizing Committee

Saša Horvat, president

University of Rijeka

Faculty of Medicine and Faculty of Health Studies

Amir Muzur

University of Rijeka

Faculty of Medicine and Faculty of Health Studies

Piotr Roszak

Nicolaus Copernicus University

Faculty of Theology

Nenad Malović

University of Zagreb

Catholic Faculty of Theology

Walter Schweidler

Catholic University of Eichstätt-Ingolstadt

Huanghao Feng

Suzhou University of Technology

School of Computer Science

Kristijan Lenac

University of Rijeka

Faculty of Engineering

Gordana Šimunković

University of Rijeka

Faculty of Medicine

Franjo Mijatović

University of Rijeka

Faculty of Medicine

Tanja Horvat

Centre for Autism Rijeka

Veronika Nela Gašpar

University of Zagreb

Catholic Faculty of Theology

Branko Murić

University of Zagreb

Catholic Faculty of Theology

Tamara Crnković

University of Rijeka Foundation

Goran Mikulić

University of Rijeka Foundation

Scientific Committee

Josip Španjol

University of Rijeka

Faculty of Medicine

Daniela Malnar

University of Rijeka

Faculty of Health Studies

Dariusz Kotecki

Nicolaus Copernicus University

Faculty of Theology

Mario Cifrak

University of Zagreb

Catholic Faculty of Theology

Tomasz Huzarek

Nicolaus Copernicus University

Faculty of Theology



SVEUČILIŠTE
U ZAGREBU

KATOLIČKI
BOGOSLOVNI
FAKULTET



We gratefully acknowledge the financial and organizational support that has made this gathering possible, provided by the Faculty of Medicine and the Faculty of Health Studies of the University of Rijeka, the University of Rijeka, the Catholic Faculty of Theology of the University of Zagreb, the Faculty of Theology of Nicolaus Copernicus University in Toruń, the University of Rijeka Foundation, the Archdiocese of Rijeka, the Metropolitan Pastoral Institute of the Archdiocese of Rijeka, the Primorje-Gorski Kotar County, the ECLAS project *Science and Faith in Healthcare: Toward a Holistic Approach* (funded by the John Templeton Foundation), and the project *Socially Assistive Robots and Autism – Perspectives and Challenges* (funded by the European Union – NextGenerationEU).

We extend our special thanks to the Hermann und Marianne Straniak-Stiftung for its significant financial support of this year's edition of the Rijeka Scientific Bridges.



EQUIPPING
CHRISTIAN
LEADERSHIP in an
AGE of
SCIENCE





**UNIWERSYTET
MIKOŁAJA KOPERNIKA
W TORUNIU**

Wydział Teologiczny



**SVEUČILIŠTE
U ZAGREBU**

**KATOLIČKI
BOGOSLOVNI
FAKULTET**



**Zaklada
Sveučilišta
u Rijeci**



9 789538 447112